

QY 67 -CAAAPAPF-----RLLW----- 79
| | | | |
Db 122 DCVPCPPGHFSPGNNOACKPWTNCTLSGKOTRHPASDSDAVCEDRSLATLLWETORPT 181
| | | | |
QY 80 -----RSCDKHTHTCCPCPAPAEAGAPSVFLFPKPKDT 112
| | | | |
Db 182 FRPTTVQSTTVWPTSELPTLVEPRSCDKHTHTCCPAPAEAGAPSVFLFPKPKDT 241
| | | | |
QY 113 LMSRTEPVTVCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 172
| | | | |
Db 242 LMSRTEPVTVCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 301
| | | | |
QY 173 QDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSREEMTKNOVSLTCLVK 232
| | | | |
Db 302 QDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSREEMTKNOVSLTCLVK 361
| | | | |
QY 233 GFYPDSIAEWESNGQPENNYKTTTPPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSYMHE 292
| | | | |
Db 362 GFYPDSIAEWESNGQPENNYKTTTPPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSYMHE 421
| | | | |
QY 293 ALHNHYTQKSLSLSPGK 309
| | | | |
Db 422 ALHNHYTQKSLSLSPGK 438
| | | | |

RESULT 2

US-08-494-574-11
; Sequence 11, Application US/08494574
; Patent No. 5783665
; GENERAL INFORMATION:
; APPLICANT: Baum, Peter
; APPLICANT: Goodwin, Ray
; APPLICANT: Fanslow, William
; APPLICANT: Gayle, Richard
; TITLE OF INVENTION: No. 5783665el Cytokine Which is a Ligand for
; TITLE OF INVENTION: OX40
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,574
; FILING DATE: 22-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,827
; FILING DATE: 23-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2806
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0730
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-494-574-11

Query Match

74.8%; Score 1255; DB 1; Length 438;

Best Local Similarity 65.8%; Pred. No. 4.3e-111;
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;
QY 35 PCSRG-SSWADLDKCMDCAS-----CRARP-----HSDFLG- 66
| | | | |
Db 62 PCETGFYNEAVNYDTCKQCTQCNRHSGSELKQNCPTPTQDVTVCRCRPGTQPRQDSGYKLG 121
| | | | |
QY 67 -CAAAPAPF-----RLLW----- 79
| | | | |
Db 122 DCVPCPPGHFSPGNNOACKPWTNCTLSGKOTRHPASDSDAVCEDRSLATLLWETORPT 181
| | | | |
QY 80 -----RSCDKHTHTCCPCPAPAEAGAPSVFLFPKPKDT 112
| | | | |
Db 182 FRPTTVQSTTVWPTSELPTLVEPRSCDKHTHTCCPAPAEAGAPSVFLFPKPKDT 241
| | | | |
QY 113 LMSRTEPVTVCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 172
| | | | |
Db 242 LMSRTEPVTVCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLH 301
| | | | |
QY 173 QDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSREEMTKNOVSLTCLVK 232
| | | | |
Db 302 QDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSREEMTKNOVSLTCLVK 361
| | | | |
QY 233 GFYPDSIAEWESNGQPENNYKTTTPPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSYMHE 292
| | | | |
Db 362 GFYPDSIAEWESNGQPENNYKTTTPPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSYMHE 421
| | | | |
QY 293 ALHNHYTQKSLSLSPGK 309
| | | | |
Db 422 ALHNHYTQKSLSLSPGK 438
| | | | |

RESULT 3
US-09-181-706-8
; Sequence 8, Application US/09181706
; Patent No. 6130068
; GENERAL INFORMATION:
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,
; APPLICANT: Robert F. DuBose, Richard S. Johnson
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/181,706
; FILING DATE: October 28, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/958,598 (converted to a
; APPLICATION NUMBER: Provisional, see below)
; FILING DATE: October 28, 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: --to be assigned-- (USN 08/958,598
; APPLICATION NUMBER: conversion to Provisional application)
; FILING DATE: October 26, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631-A
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-181-706-8

Query Match 74.5%; Score 1250; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.3e-110;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 RSCDKTHICPPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
DB 34 RSCDKTHICPPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 93
QY 140 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 199
DB 94 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 153
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 259
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 213
QY 260 LDSGSGFFLYSKLTVDKSRWQOGNPFSCSVMEALHNNHYTQKLSLSLSPGK 309
DB 214 LDSGSGFFLYSKLTVDKSRWQOGNPFSCSVMEALHNNHYTQKLSLSLSPGK 263

RESULT 4

US-09-458-791-8
Sequence 8, Application US/09458791
Patent No. 6174689

GENERAL INFORMATION:
APPLICANT: Spriggs, Melanie
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry
STREET: 51 University St.
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/Windows 95
SOFTWARE: Word for Windows 95, 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/458,791
FILING DATE: 10-Dec-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,598
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-458-791-8

Query Match 74.5%; Score 1250; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.3e-110;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 RSCDKTHICPPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
DB 34 RSCDKTHICPPCPAPEAEGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 93
QY 140 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 199
DB 94 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTIS 153
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 259
DB 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 213
QY 260 LDSGSGFFLYSKLTVDKSRWQOGNPFSCSVMEALHNNHYTQKLSLSLSPGK 309
DB 214 LDSGSGFFLYSKLTVDKSRWQOGNPFSCSVMEALHNNHYTQKLSLSLSPGK 263

RESULT 5

US-09-459-066-8
Sequence 8, Application US/09459066
Patent No. 6187909

GENERAL INFORMATION:
APPLICANT: Spriggs, Melanie
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry
STREET: 51 University St.
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/Windows 95
SOFTWARE: Word for Windows 95, 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/459,066
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,598
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-459-066-8

Query Match 74.5%; Score 1250; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.3e-110;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 RSCDKTHTCCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
Db 34 RSCDKTHTCCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 93
Qy 140 YVDGVEVHNKTKPREQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 199
Db 94 YVDGVEVHNKTKPREQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 153
Qy 200 KAGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPV 259
Db 154 KAGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPV 213
Qy 260 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 309
Db 214 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 263

RESULT 6

US-09-180-100-22
; Sequence 22, Application US/09180100
; Patent No. 6306395
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, No. 6306395iio
; APPLICANT: NAGATA, Shigekazu
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE
; FILE REFERENCE: 1110-207p
; CURRENT APPLICATION NUMBER: US/09/180,100
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: PCT/JP97/01502
; EARLIER FILING DATE: 1997-05-01
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-180-100-22

Query Match 74.3%; Score 1247; DB 4; Length 376;
Best Local Similarity 65.4%; Pred. No. 2e-110;
Matches 249; Conservative 10; Mismatches 32; Indels 90; Gaps 8;

Qy 14 VLGLWAL---LRSA-----GEQ-----APGTAPCSRGS 40
Db 1 MGLWTLLPLVLTSAQNLEGLHDGQFCHKPCPGERRKARDCTVNGDEPDVPCQEGK 60
Qy 41 SWSADL---DKMDC-----ASCRAP-----HSDFCIGC 67
Db 61 EYTDKAHFSKRCRCRLCDEGHGHEVEINCTRTQNTKCRCKPNFNCNSTVCEHCDPCTKC 120
Qy 68 -----AAAPAPRLLWRSCKDTHCTPCCPAPAEAGAPSVFLPPK 108
Db 121 EHGIIKCTLTNTKCKEGRSNEP-----KSCDKTHTCCPAPPELLGGPSVFLPPK 175
Qy 109 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVVSVL 168
Db 176 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVVSVL 235
Qy 169 TVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLT 228
Db 236 TVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLT 295
Qy 229 CLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSLKLTVDKSRWQGNVFCSC 288
Db 296 CLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSLKLTVDKSRWQGNVFCSC 355
Qy 289 VMHEALHNYTKQSLSLSPGK 309
Db 356 VMHEALHNYTKQSLSLSPGK 376

RESULT 7

US-09-180-100-11
; Sequence 11, Application US/09180100
; Patent No. 6306395
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, No. 6306395iio
; APPLICANT: NAGATA, Shigekazu
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE
; FILE REFERENCE: 1110-207p
; CURRENT APPLICATION NUMBER: US/09/180,100
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: PCT/JP97/01502
; EARLIER FILING DATE: 1997-05-01
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-180-100-11

Query Match 74.1%; Score 1244; DB 4; Length 360;
Best Local Similarity 71.3%; Pred. No. 3.6e-110;
Matches 241; Conservative 10; Mismatches 29; Indels 58; Gaps 6;

Qy 24 SVAGEQAPGTAPCSRGSWSADL---DKMDC-----AAAPAPRLLWRSCKDTHCTPCPC 59
Db 29 TVNGDE-PCVPCQEGKEYTDKAHFSKRCRCRLCDEGHGHEVEINCTRTQNTKCRCKPN 87
Qy 60 -----HSDFCIGC-----KSCDKTHTCCP 91
Db 88 FFCNSTVCEHCDPCTKCEHGIKCTLTNTKCKEGRSNEP-----KSCDKTHTCCP 142
Qy 92 PAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNK 151
Db 143 PAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNK 202
Qy 152 KPREEQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPOVY 211
Db 203 KPREEQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPOVY 262
Qy 212 TLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSK 271
Db 263 TLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSK 322
Qy 272 LTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 309
Db 323 LTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPGK 360

RESULT 8

PCT-US95-03866-12
; Sequence 12, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: Cytomed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

Sequence 3, Application US/08784512
Patent No. 5872209
GENERAL INFORMATION:
APPLICANT: BARTNIK, Eckart
APPLICANT: EIDENMUELLER, Bernd
APPLICANT: BUETTNER, Frank
APPLICANT: CATERSON, Bruce
APPLICANT: HUGHES, Clare
TITLE OF INVENTION: An artificial recombinant substrate (TAGG 1)
TITLE OF INVENTION: and native aggrecan to study the proteolytic activity of
TITLE OF INVENTION: "Aggrecanase" in cell culture systems
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,512
FILING DATE: 17-JAN-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 96100682.2
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 18748/311
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..396
US-08-784-512-3

Query Match 73.5%; Score 1233.5; DB 2; Length 396;
Best Local Similarity 83.0%; Pred. No. 4.1e-109;
Matches 235; Conservative 8; Mismatches 27; Indels 13; Gaps 3;
QY 27 GEQAGTAPCSRSSWSADLDKCMDCASCARPHSDFCGLCAAPAPFLLWRSCDKTH 86
DB 127 GFPTPLGLPATAFTS-----EDLVVQVTVAVPGQPH----LPGGGDPEP-----KSCDKTH 173
QY 87 TCPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 146
DB 174 TCPCPAPELLGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 233
QY 147 HNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 206
DB 234 HNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 293
QY 207 EPQVYILPPSREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSF 266
DB 294 EPQVYILPPSRDELTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSF 353
QY 267 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 309
DB 354 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 396

RESULT 11
US-09-176-228-3
Sequence 3, Application US/09176228
Patent No. 6180334
GENERAL INFORMATION:
APPLICANT: BARTNIK, Eckart
APPLICANT: EIDENMUELLER, Bernd
APPLICANT: BUETTNER, Frank
APPLICANT: CATERSON, Bruce
APPLICANT: HUGHES, Clare
TITLE OF INVENTION: An artificial recombinant substrate (rAGG 1)
TITLE OF INVENTION: and native aggregan to study the proteolytic activity of
TITLE OF INVENTION: "Aggrecanase" in cell culture systems
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/176,228
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/784,512
FILING DATE: 17-JAN-1997
APPLICATION NUMBER: EP 96100682.2
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 18748/311
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..396
US-09-176-228-3
Query Match 73.5%; Score 1233.5; DB 4; Length 396;
Best Local Similarity 83.0%; Pred. No. 4.1e-109;
Matches 235; Conservative 8; Mismatches 27; Indels 13; Gaps 3;
QY 27 GEQAGTAPCSRSSWSADLDKCMDCASCARPHSDFCGLCAAPAPFLLWRSCDKTH 86
DB 127 GFPTPLGLPATAFTS-----EDLVVQVTVAVPGQPH----LPGGGDPEP-----KSCDKTH 173
QY 87 TCPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 146
DB 174 TCPCPAPELLGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 233
QY 147 HNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 206
DB 234 HNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 293
QY 207 EPQVYILPPSREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSF 266
DB 294 EPQVYILPPSRDELTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSF 353
QY 267 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 309
DB 354 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 396

RESULT 12
US-08-887-352B-14
Sequence 14, Application US/08887352B
Patent No. 5994511
GENERAL INFORMATION:
APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
TITLE OF INVENTION: Improving Polypeptides
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,352B
FILING DATE: 03-Jul-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1123
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-887-352B-14

Query Match 73.3%; Score 1230; DB 2; Length 451;
Best Local Similarity 98.3%; Pred. No. 1.le-108;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHCPCPAPEAGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
DB 222 KSCDKTHCPCPAPELGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281
QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
DB 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYGPPSDIAVWESNGQPENNYKTTPPV 259
DB 342 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYGPPSDIAVWESNGQPENNYKTTPPV 401
QY 260 LDSGSEFLYSLKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPCK 309
DB 402 LDSGSEFLYSLKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPCK 451

RESULT 13
US-08-887-352B-16
Sequence 16, Application US/08887352B
Patent No. 5994511
GENERAL INFORMATION:
APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
TITLE OF INVENTION: Improving Polypeptides
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,352B
FILING DATE: 03-Jul-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1123
TELEPHONE: 650/225-1489

TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-887-352B-16

Query Match 73.3%; Score 1230; DB 2; Length 451;
Best Local Similarity 98.3%; Pred. No. 1.le-108;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHCPCPAPEAGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
DB 222 KSCDKTHCPCPAPELGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281
QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
DB 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341
QY 200 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYGPPSDIAVWESNGQPENNYKTTPPV 259
DB 342 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYGPPSDIAVWESNGQPENNYKTTPPV 401
QY 260 LDSGSEFLYSLKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPCK 309
DB 402 LDSGSEFLYSLKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPCK 451

RESULT 14
US-08-887-352B-18
Sequence 18, Application US/08887352B
Patent No. 5994511
GENERAL INFORMATION:

APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
TITLE OF INVENTION: Improving Polypeptides
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,352B
FILING DATE: 03-Jul-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1123
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-887-352B-18

Query Match 73.3%; Score 1230; DB 2; Length 451;
Best Local Similarity 98.3%; Pred. No. 1.le-108;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHCPCPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
:|||||
Db 222 KSCDKTHCPCPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281
:|||||
QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
:|||||
Db 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341
:|||||
QY 200 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 259
:|||||
Db 342 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 401
:|||||
QY 260 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCSVNHEALHNNHYTKQSLSPGK 309
:|||||
Db 402 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCSVNHEALHNNHYTKQSLSPGK 451
:|||||

RESULT 15
US-08-466-151-65
: Sequence 65, Application US/08466151
: Patent No. 6037453
: GENERAL INFORMATION:
: APPLICANT: Jardieu, Paula M.
: APPLICANT: Presta, Leonard G.
: TITLE OF INVENTION: Immunoglobulin Variants
: NUMBER OF SEQUENCES: 65
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Winpatin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/466,151
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/466163
: FILING DATE: 06-Jun-1995
: APPLICATION NUMBER: 08/405617
: FILING DATE: 15-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/185899
: FILING DATE: 26-JAN-1994
: PRIOR APPLICATION DATA: 07/879495
: APPLICATION NUMBER: 07/879495
: FILING DATE: 07-MAY-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/744768
: FILING DATE: 14-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Svoboda, Craig G.
: REGISTRATION NUMBER: 39,044
: REFERENCE/DOCKET NUMBER: P0718P2C1D1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650/225-1489
: TELEFAX: 650/952-9881
: INFORMATION FOR SEQ ID NO: 65:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 451 amino acids
: TYPE: Amino Acid
: TOPOLOGY: Linear
: US-08-466-151-65

Query Match 73.3%; Score 1230; DB 3; Length 451;

Best Local Similarity 98.3%; Pred. No. 1.1e-108;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 80 RSCDKTHCPCPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
:|||||
Db 222 KSCDKTHCPCPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281
:|||||
QY 140 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
:|||||
Db 282 YVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341
:|||||
QY 200 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 259
:|||||
Db 342 KAKGQPREPOVYITLPPSRREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 401
:|||||
QY 260 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCSVNHEALHNNHYTKQSLSPGK 309
:|||||
Db 402 LDSGSEFFLYSKLTVDKSRWQOQGNVFSCSVNHEALHNNHYTKQSLSPGK 451
:|||||

Search completed: March 11, 2002, 15:13:12
Job time: 36 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:13:14 ; Search time 26.47 Seconds
(without alignments)
239.740 Million cell updates/sec

Title: US-09-742-454A-7_COPY_28_309

Perfect score: 1551

Sequence: 1 EQAPGAPCSRSGSSWADLD.....MHEALHHYTKSLSPGK 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTRUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1255	80.9	438	1	US-08-097-827-11
2	1255	80.9	438	1	US-08-494-574-11
3	1250	80.6	660	4	US-09-181-706-8
4	1250	80.6	660	4	US-09-458-791-8
5	1250	80.6	660	4	US-09-459-066-8
6	1243	80.1	424	5	PCT-US95-03866-12
7	1243	80.1	424	5	PCT-US95-03866-14
8	1241.5	80.0	360	4	US-09-180-100-11
9	1241.5	80.0	376	4	US-09-180-100-22
10	1231.5	79.4	396	2	US-08-784-512-3
11	1231.5	79.4	396	4	US-09-176-228-3
12	1230	79.3	451	2	US-08-887-352B-14
13	1230	79.3	451	2	US-08-887-352B-16
14	1230	79.3	451	2	US-08-887-352B-18
15	1230	79.3	451	3	US-08-466-151-65
16	1230	79.3	451	4	US-09-109-207C-14
17	1230	79.3	451	4	US-09-109-207C-16
18	1230	79.3	451	4	US-09-109-207C-18
19	1230	79.3	451	4	US-09-282-505-2
20	1230	79.3	451	4	US-09-054-255-2
21	1230	79.3	451	4	US-09-296-005-14
22	1230	79.3	451	4	US-09-296-005-16
23	1230	79.3	451	4	US-09-296-005-18
24	1230	79.3	452	3	US-09-027-449-71
25	1230	79.3	452	4	US-09-026-985-71
26	1230	79.3	453	3	US-08-466-151-8
27	1230	79.3	454	2	US-07-934-373C-22

28	1230	79.3	454	3	US-08-437-642B-22	Sequence 22, Appl
29	1230	79.3	454	5	PCT-US93-07832-22	Sequence 22, Appl
30	1230	79.3	467	4	US-09-049-672A-8	Sequence 8, Appl
31	1230	79.3	473	4	US-09-049-672A-4	Sequence 4, Appl
32	1225.5	79.0	442	5	PCT-US96-10043-9	Sequence 9, Appl
33	1224	78.9	232	2	US-08-595-043A-50	Sequence 50, Appl
34	1224	78.9	331	4	US-09-178-869-2	Sequence 2, Appl
35	1224	78.9	371	1	US-08-236-311-7	Sequence 7, Appl
36	1224	78.9	371	3	US-08-457-918-7	Sequence 7, Appl
37	1224	78.9	437	5	PCT-US96-10043-11	Sequence 11, Appl
38	1224	78.9	446	3	US-08-397-411-7	Sequence 7, Appl
39	1224	78.9	449	1	US-08-458-516-13	Sequence 13, Appl
40	1224	78.9	459	1	US-08-157-101A-7	Sequence 7, Appl
41	1224	78.9	476	2	US-08-378-939-10	Sequence 10, Appl
42	1224	78.9	476	3	US-08-487-550-4	Sequence 4, Appl
43	1224	78.9	476	3	US-08-487-550-12	Sequence 12, Appl
44	1224	78.9	478	3	US-08-487-550-8	Sequence 8, Appl
45	1223	78.9	232	3	US-08-996-139-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-097-827-11
; Sequence 11, Application US/08097827
; GENERAL INFORMATION:
; APPLICANT: Baum, Peter
; Goodwin, Ray
; Fanslow, William
; Gayle, Richard
; TITLE OF INVENTION: Novel Cytokine which is a Ligand for
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/097,827
; FILING DATE: 23-Jul-1993
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2806
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0730
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-097-827-11

Query Match 80.9%; Score 1255; DB 1; Length 438;
Best Local Similarity 65.8%; Pred. No. 7.8e-113;
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;
QY 8 PCSRG-SWSADLDCMDCAS-----CRARP-----HSDFCLG- 39
||| : : : : : ||| : : : : :
DB 62 PCETGYNEAVNYDCKQCTQCNRHSGSELKQNCIPTQDTVCRCRGTPTQDRODSGYKLG 121

```
QY 40 -CAAAPAPF-----RLLW----- 52
Db 122 DCVPCPGHSPGNNQACKPWTNCTLSGKQTRHPASDSDAVCEDRSLATLLWETORPT 181
QY 53 -----RSCDKHTHTCCPPAPAEAGAPSVFLFPPKPKDT 85
Db 182 FRPTTVQSTTVWPTSELPSTPTLVEPRSCDKHTHTCCPPAPAEAGAPSVFLFPPKPKDT 241
QY 86 LMSRTPETVTCVVVDVSHEDPEVKFNNYVVGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 145
Db 242 LMSRTPETVTCVVVDVSHEDPEVKFNNYVVGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 301
QY 146 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREEMTKNOVSLTCLVK 205
Db 302 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 361
QY 206 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSYMHE 265
Db 362 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSYMHE 421
QY 266 ALHNHYTQKSLSLSPGK 282
Db 422 ALHNHYTQKSLSLSPGK 438

RESULT 2
US-08-494-574-11
; Sequence 11, Application US/08494574
; Patent No. 5783665
; GENERAL INFORMATION:
; APPLICANT: Baum, Peter
; APPLICANT: Goodwin, Ray
; APPLICANT: Fauslow, William
; APPLICANT: Gayle, Richard
; TITLE OF INVENTION: No. 5783665el Cytokine which is a Ligand for
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,574
; FILING DATE: 22-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,827
; FILING DATE: 23-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2806
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0730
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-494-574-11
```

Query Match 80.9%; Score 1255; DB 1; Length 438;

```
Best Local Similarity 65.8%; Pred. No. 7.8e-113;
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;

QY 8 PCSRG-SSWADLDKCMDCAS-----CRARP-----HSDFLG- 39
Db 62 PCBTGFYNEAVNYDTCKQCTQCNRHSGSELKQNCPTPTQDTCRCRPGTQPRODSGYKLG 121
QY 40 -CAAAPAPF-----RLLW----- 52
Db 122 DCVPCPGHSPGNNQACKPWTNCTLSGKQTRHPASDSDAVCEDRSLATLLWETORPT 181
QY 53 -----RSCDKHTHTCCPPAPAEAGAPSVFLFPPKPKDT 85
Db 182 FRPTTVQSTTVWPTSELPSTPTLVEPRSCDKHTHTCCPPAPAEAGAPSVFLFPPKPKDT 241
QY 86 LMSRTPETVTCVVVDVSHEDPEVKFNNYVVGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 145
Db 242 LMSRTPETVTCVVVDVSHEDPEVKFNNYVVGVEVHNAKTKPREEQYNSTYRVVSVLTVLH 301
QY 146 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREEMTKNOVSLTCLVK 205
Db 302 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 361
QY 206 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSYMHE 265
Db 362 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSYMHE 421
QY 266 ALHNHYTQKSLSLSPGK 282
Db 422 ALHNHYTQKSLSLSPGK 438

RESULT 3
US-09-181-706-8
; Sequence 8, Application US/09181706
; Patent No. 6130068
; GENERAL INFORMATION:
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,
; APPLICANT: Robert F. DuBose, Richard S. Johnson
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/181,706
; FILING DATE: October 28, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/958,598 (converted to a
; APPLICATION NUMBER: Provisional, see below)
; FILING DATE: October 28, 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: --to be assigned-- (USN 08/958,598
; APPLICATION NUMBER: conversion to Provisional application)
; FILING DATE: October 26, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631-A
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-181-706-8

Query Match 80.6%; Score 1250; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 4.2e-112;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEGAPSVFLPPPKDFTLMSRTPEVTCVVVDVSHEDPEVKFNW 112
Db 34 RSCDKTHTCPPCPAPEGAPSVFLPPPKDFTLMSRTPEVTCVVVDVSHEDPEVKFNW 93
QY 113 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172
Db 94 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 153
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 232
Db 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 213
QY 233 LSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 282
Db 214 LSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 263

RESULT 4

US-09-458-791-8
Sequence 8, Application US/09458791
Patent No. 6174689
GENERAL INFORMATION:
APPLICANT: Spriggs, Melanie
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
RECEPTOR DNA AND POLYPEPTIDES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry
STREET: 51 University St.
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/Windows 95
SOFTWARE: Word for Windows 95, 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/458,791
FILING DATE: 10-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,598
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-458-791-8

Query Match 80.6%; Score 1250; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 4.2e-112;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEGAPSVFLPPPKDFTLMSRTPEVTCVVVDVSHEDPEVKFNW 112
Db 34 RSCDKTHTCPPCPAPEGAPSVFLPPPKDFTLMSRTPEVTCVVVDVSHEDPEVKFNW 93
QY 113 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172
Db 94 YVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 153
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 232
Db 154 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 213
QY 233 LSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 282
Db 214 LSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 263

RESULT 5

US-09-459-066-8
Sequence 8, Application US/09459066
Patent No. 6187909
GENERAL INFORMATION:
APPLICANT: Spriggs, Melanie
TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
RECEPTOR DNA AND POLYPEPTIDES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry
STREET: 51 University St.
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/Windows 95
SOFTWARE: Word for Windows 95, 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/459,066
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,598
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-459-066-8

Query Match 80.6%; Score 1250; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 4.2e-112;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	53	RSCDKHTCPCPADPAEAGAPSVFLPPPKDPTLMTSRTPVTCVVVDVSHEDPEVKFNN	112
Db	34	RSCDKHTCPCPADPAEAGAPSVFLPPPKDPTLMTSRTPVTCVVVDVSHEDPEVKFNN	93
Qy	113	YVDGVEVHNAKTRPREQYNSTIRRVYSVLTFLVLDHQLWLNCKEYCKVSNKALPAPTEKTLIS	172
Db	94	YVDGVEVHNAKTRPREQYNSTIRRVYSVLTFLVLDHQLWLNCKEYCKVSNKALPAPTEKTLIS	153
Qy	173	KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV	232
Db	154	KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV	213
Qy	233	LDSDGFFFLYSKLTVDKSRWQQGNVSCSVMHEALNHHYTKQSLSLSPGK	282
Db	214	LDSDGFFFLYSKLTVDKSRWQQGNVSCSVMHEALNHHYTKQSLSLSPGK	263

```

RESULT      6
PCT-US95-03866-12
; Sequence 12, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: Cytomed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CytoMed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-03866-12

```

```

Query Match      80.1%; Score 1243; DB 5; Length 424;
Best Local Similarity 84.6%; Pred. No. 1.le-111;
Matches 237; Conservative 6; Mismatches 25; Indels 12; Gaps 2;

Qy 3 APTAPCGRSSGADLDKMCDCASCRARPHSDCLGCAAPAPFRLLWRSCDKTHTCP 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 157 ASETSDCWSSSTLSPKSDS-----RVSYTKPMLPEVAADPEP-----KSCDKTHTCP 204

Qy 63 PCPAPEAEGAPSVFLPPPKPKDTLMI SRTPEVTCVVVDVSHEDPEVKFNWYDVGVEVHNA 122
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 205 PCPAPELLGGPSVFLPPPKPKDTLMI SRTPEVTCVVVDVSHEDPEVKFNWYDVGVEVHNA 264

```

Qy	123	KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREP	182
Db	265	KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREP	324
Qy	183	VYTLPPSREEMTKNQVSLTCLLVGFYPSDIAVWESNGQPENNYKTTTPVLDSDGSFFLY	242
Db	325	VYTLPPSRDELTKNQVSLTCLLVGFYPSDIAVWESNGQPENNYKTTTPVLDSDGSFFLY	384
Qy	243	SKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK	282
Db	385	SKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK	424

```

7
RESULT
PCT-US95-03866-14
: Sequence 14, Application PC/TUS9503866
: GENERAL INFORMATION:
: APPLICANT: Cytomed, Inc. (all states except US)
: APPLICANT: Nocka, Karl (US only)
: APPLICANT: Lobell, Robert B (US only)
: TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
: TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Neave
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/03866
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/220,379
: FILING DATE: 28-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Haley Jr, James F
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: CytoMed/2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-596-9000
: TELEFAX: 212-596-9090
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 424 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-03866-14

```

Query Match	80.1%	Score 1243;	DB 5;	Length 424;
Best Local Similarity	84.6%	Pred. No. 1.1e-11;		
Matches 237;	Conservative	6;	Mismatches 25;	
			Indels 12;	Gaps 2;

QY 123 KTKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGPREPQ 182

Db 265 KTKPREEQNSTYRVSVLTVLHQLDNLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQ 324
QY 183 VYTLPSREMTKNQVSLCLVKGFPYPSDIAVEWESNGQPNENYKTTTPVLDSDGSFFLY 242
Db 325 VYTLPPSRDELTKNQVSLCLVKGFPYPSDIAVEWESNGQPNENYKTTTPVLDSDGSFFLY 384
QY 243 SKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 282
Db 385 SKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 424

RESULT 8

US-09-180-100-11
; Sequence 11, Application US/09180100
; Patent No. 6306395
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, No. 630639510
; APPLICANT: NAGATA, Shigekazu
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE
; FILE REFERENCE: 1110-207P
; CURRENT APPLICATION NUMBER: US/09/180,100
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: PCT/JP97/01502
; EARLIER FILING DATE: 1997-05-01
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-180-100-11

Query Match 80.0%; Score 1241.5; DB 4; Length 360;
Best Local Similarity 72.2%; Pred. No. 1.2e-111;
Matches 239; Conservative 7; Mismatches 28; Indels 57; Gaps 5;

QY 4 PGTAPCSRGSSWSADL---DKMDC-----ASCRARP----- 32
Db 35 PDCVPCQEGKEYTDKAHFSSKRCRLCDEGHGLEVEINCTRTQNTKCRCKPNFFCNSTV 94
QY 33 --HSDFCCLG-----AAAPPAPFRLLRSCDKTHTCPCPAPEARG 71
Db 95 CEHCDPCTKCEHGIKECTLTNTKCKEGRSNEP-----KSCDKTHTCPCPAPELIG 149
QY 72 APSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPREEQY 131
Db 150 GPSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPREEQY 209
QY 132 NSTYRVSVLTVLHODLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRE 191
Db 210 NSTYRVSVLTVLHODLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRD 269
QY 192 EMTKNQVSLCLVKGFPYPSDIAVEWESNGQPNENYKTTTPVLDSDGSFFLYSKLTVDKSR 251
Db 270 ELTKNQVSLCLVKGFPYPSDIAVEWESNGQPNENYKTTTPVLDSDGSFFLYSKLTVDKSR 329
QY 252 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 282
Db 330 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 360

RESULT 9

US-09-180-100-22
; Sequence 22, Application US/09180100
; Patent No. 6306395
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, No. 630639510
; APPLICANT: NAGATA, Shigekazu
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE
; FILE REFERENCE: 1110-207P
; CURRENT APPLICATION NUMBER: US/09/180,100
; CURRENT FILING DATE: 1998-11-02

; EARLIER APPLICATION NUMBER: PCT/JP97/01502
; EARLIER FILING DATE: 1997-05-01
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-180-100-22

Query Match 80.0%; Score 1241.5; DB 4; Length 376;
Best Local Similarity 72.2%; Pred. No. 1.2e-111;
Matches 239; Conservative 7; Mismatches 28; Indels 57; Gaps 5;

QY 4 PGTAPCSRGSSWSADL---DKMDC-----ASCRARP----- 32
Db 51 PDCVPCQEGKEYTDKAHFSSKRCRLCDEGHGLEVEINCTRTQNTKCRCKPNFFCNSTV 110
QY 33 --HSDFCCLG-----AAAPPAPFRLLRSCDKTHTCPCPAPEARG 71
Db 111 CEHCDPCTKCEHGIKECTLTNTKCKEGRSNEP-----KSCDKTHTCPCPAPELIG 165
QY 72 APSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPREEQY 131
Db 166 GPSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPREEQY 225
QY 132 NSTYRVSVLTVLHODLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRE 191
Db 226 NSTYRVSVLTVLHODLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRD 285
QY 192 EMTKNQVSLCLVKGFPYPSDIAVEWESNGQPNENYKTTTPVLDSDGSFFLYSKLTVDKSR 251
Db 286 ELTKNQVSLCLVKGFPYPSDIAVEWESNGQPNENYKTTTPVLDSDGSFFLYSKLTVDKSR 345
QY 252 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 282
Db 346 WQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 376

RESULT 10

US-08-784-512-3
; Sequence 3, Application US/08784512
; Patent No. 5872209
; GENERAL INFORMATION:
; APPLICANT: BARTNIK, Eckart
; APPLICANT: EIDENMUELLER, Bernd
; APPLICANT: BUETTNER, Frank
; APPLICANT: CATERSON, Bruce
; APPLICANT: HUGHES, Clare
; TITLE OF INVENTION: An artificial recombinant substrate (rAGG 1)
; TITLE OF INVENTION: and native aggreca to study the proteolytic activity of
; TITLE OF INVENTION: "Aggrecanase" in cell culture systems
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,512
; FILING DATE: 17-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 96100682.2
; FILING DATE: 18-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.

REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 18748/311
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:

NAME/KEY: Protein
LOCATION: 1..396
US-08-784-512-3

Query Match 79.4%; Score 1231.5; DB 2; Length 396;

Best Local Similarity 83.9%; Pred. No. 1.2e-110; Mismatches 24; Indels 13; Gaps 3;

Matches 234; Conservative 8; Mismatches 24; Indels 13; Gaps 3;

QY 4 PGAPCRSGSSWADLDKCMDCASCARPHSDFCGCAAAAPPAPFRLLRSCDKTHTCPP 63
DB 131 PGLGPATAFTS----EDLVQVTVAPGQPH----LPGGDPPEP-----KSCDKTHTCPP 177
QY 64 CPAPEAGASVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 123
DB 178 CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 237
QY 124 TKPREEOYNSTYRVSVTLVTHQDMLNGKEYCKVSKNKAAPAPLEKTSIRAKGQPREPOV 183
DB 238 TKPREEOYNSTYRVSVTLVTHQDMLNGKEYCKVSKNKAAPAPLEKTSIRAKGQPREPOV 297
QY 184 YTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFELY 243
DB 298 YTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFELY 357
QY 244 KLTVDKSRWQGNVFCSCVMHEALHNYHNTOKSLSPGK 282
DB 358 KLTVDKSRWQGNVFCSCVMHEALHNYHNTOKSLSPGK 396

RESULT 11

US-09-176-228-3

Sequence 3, Application US/09176228

Patent No. 6180334

GENERAL INFORMATION:

APPLICANT: BARTNIK, Eckart
APPLICANT: EIDENMUELLER, Bernd
APPLICANT: BUETTNER, Frank
APPLICANT: CATERSON, Bruce
APPLICANT: HUGHES, Clare

TITLE OF INVENTION: An artificial recombinant substrate (RAGG 1)
TITLE OF INVENTION: and native aggregan to study the proteolytic activity of

TITLE OF INVENTION: "Aggreganase" in cell culture systems

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: Foley & Lardner

STREET: Suite 500, 3000 K Street, N.W.

CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/176,228

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/784,512
FILING DATE: 17-JAN-1997
APPLICATION NUMBER: EP 96100682.2
FILING DATE: 18-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: GRANADOS, Patricia D.

REGISTRATION NUMBER: 33,683

REFERENCE/DOCKET NUMBER: 18748/311

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 396 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Protein

LOCATION: 1..396

US-09-176-228-3

Query Match 79.4%; Score 1231.5; DB 4; Length 396;

Best Local Similarity 83.9%; Pred. No. 1.2e-110; Mismatches 24; Indels 13; Gaps 3;

Matches 234; Conservative 8; Mismatches 24; Indels 13; Gaps 3;

QY 4 PGAPCRSGSSWADLDKCMDCASCARPHSDFCGCAAAAPPAPFRLLRSCDKTHTCPP 63
DB 131 PGLGPATAFTS----EDLVQVTVAPGQPH----LPGGDPPEP-----KSCDKTHTCPP 177
QY 64 CPAPEAGASVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 123
DB 178 CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 237
QY 124 TKPREEOYNSTYRVSVTLVTHQDMLNGKEYCKVSKNKAAPAPLEKTSIRAKGQPREPOV 183
DB 238 TKPREEOYNSTYRVSVTLVTHQDMLNGKEYCKVSKNKAAPAPLEKTSIRAKGQPREPOV 297
QY 184 YTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFELY 243
DB 298 YTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFELY 357
QY 244 KLTVDKSRWQGNVFCSCVMHEALHNYHNTOKSLSPGK 282
DB 358 KLTVDKSRWQGNVFCSCVMHEALHNYHNTOKSLSPGK 396

RESULT 12

US-08-887-352B-14

Sequence 14, Application US/08887352B

Patent No. 5994511

GENERAL INFORMATION:

APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe

TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of

TITLE OF INVENTION: Improving Polypeptides

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESS: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/887,352B

;
; FILING DATE: 03-Jul-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-887-352B-14

Query Match 79.3%; Score 1230; DB 2; Length 451;
Best Local Similarity 98.3%; Pred. No. 2.1e-110;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDKTHTCTPCPAPAEAGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 112
Db 222 KSCDKTHTCTPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 281
QY 113 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172
Db 282 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341
QY 173 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPV 232
Db 342 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPV 401
QY 233 LSDSGSFYLYSLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSLSPGK 282
Db 402 LSDSGSFYLYSLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSLSPGK 451

RESULT 13

US-08-887-352B-16
; Sequence 16, Application US/08887352B
; Patent No. 5994511
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardiou, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
; TITLE OF INVENTION: Improving Polypeptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,352B
; FILING DATE: 03-Jul-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids

;
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-887-352B-16

Query Match 79.3%; Score 1230; DB 2; Length 451;
Best Local Similarity 98.3%; Pred. No. 2.1e-110;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDKTHTCTPCPAPAEAGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 112
Db 222 KSCDKTHTCTPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 281
QY 113 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172
Db 282 YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 341
QY 173 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPV 232
Db 342 KAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPV 401
QY 233 LSDSGSFYLYSLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSLSPGK 282
Db 402 LSDSGSFYLYSLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSLSPGK 451

RESULT 14

US-08-887-352B-18
; Sequence 18, Application US/08887352B
; Patent No. 5994511
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardiou, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
; TITLE OF INVENTION: Improving Polypeptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,352B
; FILING DATE: 03-Jul-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-887-352B-18

Query Match 79.3%; Score 1230; DB 2; Length 451;
Best Local Similarity 98.3%; Pred. No. 2.1e-110;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDKTHTCTPCPAPAEAGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 112
Db 222 KSCDKTHTCTPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNW 281

QY 113 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTIS 172
|||||
Db 282 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTIS 341
|||||
QY 173 KAKQPREPOQVYTLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTPPV 232
|||||
Db 342 KAKQPREPOQVYTLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTPPV 401
|||||
QY 233 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 282
|||||
Db 402 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 451
|||||

RESULT 15
US-08-466-151-65
; Sequence 65, Application US/08466151
; Patent No. 6037453
; GENERAL INFORMATION:
; APPLICANT: Jardiou, Paula M.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,151
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/466163
; FILING DATE: 06-Jun-1995
; APPLICATION NUMBER: 08/405617
; FILING DATE: 15-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/185899
; FILING DATE: 26-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/879495
; FILING DATE: 07-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P2C1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-466-151-65

Query Match 79.38; Score 1230; DB 3; Length 451;
Best Local Similarity 98.38; Pred. No. 2.le-110;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 53 RSCDKTHTCPCPAPEAGAPSVLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112

Db 222 KSCDKTHTCPCPAPELGGPSVLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 281
|||||
QY 113 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTIS 172
|||||
Db 282 YVDGVEVHNAKTPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTIS 341
|||||
QY 173 KAKQPREPOQVYTLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTPPV 232
|||||
Db 342 KAKQPREPOQVYTLPPSREEMTKNOVSLTCLVKGIFYPSDIAVEWESNGQPENNYKTTPPV 401
|||||
QY 233 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 282
|||||
Db 402 LDSGSEFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQKLSLSPGK 451
|||||

Search completed: March 11, 2002, 15:13:15
Job time: 39 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:14 ; Search time 53.92 Seconds
(without alignments)
387.401 Million cell updates/sec

Title: US-09-742-454A-7_COPY_28_309

Perfect score: 1551

Sequence: 1 EQAPGTAPCSRGSSWEADLD.....MHEALHNHYTKSLSPGK 282

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1551	100.0	309	22	AAU03500	Human TWEAKR recept
2	1258	81.1	451	20	AAV43139	NAIL-Fc protein se
3	1255	80.9	438	16	AAR81882	Plasmid pDC406/OX4
4	1255	80.9	438	19	AAW48976	OX40/Fc mutein. C
5	1250	80.6	660	21	AAV13463	Ectromelia A39R se
6	1250	80.6	660	21	AAV13463	Ectromelia A39R se
7	1250	80.6	660	22	AAV13463	Ectromelia A39R se
8	1250	80.6	660	22	AAV13463	Ectromelia A39R se
9	1247	80.4	450	20	AAV15239	Amino acid sequenc
10	1247	80.4	453	20	AAV15239	ULBP1-Fc fusion po
11	1243	80.1	424	16	AAW14764	ULBP2-Fc fusion po
						Human soluble kit

12	1243	80.1	424	16	AAW14765	Human soluble kit
13	1241.5	80.0	376	19	AAW60037	Antigenic peptide
14	1239.5	79.9	375	18	AAW06683	Heregulin-alpha fu
15	1239	79.9	535	20	AAW17414	SYPH1-26 disintegr
16	1235.5	79.7	234	18	AAW16686	Human IgG1 FC port
17	1235	79.6	400	21	AAV15123	Porcine CTLA-4-Ig
18	1231.5	79.4	396	18	AAW18574	Aggrecanase artifi
19	1231.5	79.4	396	18	AAW18575	Aggrecanase artifi
20	1230.5	79.3	488	21	AAV97175	Human FGF-RI Extra
21	1230	79.3	232	22	AAE02642	Human immunoglobul
22	1230	79.3	292	22	AAE02646	Human immunoglobul
23	1230	79.3	292	22	AAE02647	Human IgG1 FC regi
24	1230	79.3	330	20	AAV50153	Human heavy chain
25	1230	79.3	339	22	AAE02648	Human immunoglobul
26	1230	79.3	401	18	AAW10537	Leptin 1-167/IgG1
27	1230	79.3	444	21	AAV32263	Humanised anti-CD2
28	1230	79.3	451	20	AAV50031	Human E27 anti-IgE
29	1230	79.3	451	20	AAW95659	Mus musculus anti-
30	1230	79.3	451	20	AAW95661	Mus musculus anti-
31	1230	79.3	451	20	AAW95663	Mus musculus anti-
32	1230	79.3	451	21	AAW07473	Amino acid sequenc
33	1230	79.3	451	21	AAV85201	Light chain amino
34	1230	79.3	451	22	AAV74212	E27 anti-IgE antib
35	1230	79.3	451	22	AAV74088	Anti-IgE antibody,
36	1230	79.3	451	22	AAV76948	Full length heavy
37	1230	79.3	451	22	AAV76950	Full length heavy
38	1230	79.3	451	22	AAV76952	Full length heavy
39	1230	79.3	452	19	AAW69316	Anti-IL-8 humanise
40	1230	79.3	452	20	AAV29458	Recombinant immuno
41	1230	79.3	452	21	AAV30322	Humanised anti-IL-
42	1230	79.3	452	21	AAV77666	Humanised anti-IL-
43	1230	79.3	453	14	AAV33311	Humanised MaE11 Ve
44	1230	79.3	453	20	AAV50151	Antibody F19 chime
45	1230	79.3	453	21	AAV85199	Heavy chain amino

ALIGNMENTS

RESULT 1

AAU03500

ID AAU03500 standard; Protein; 309 AA.

AC AAU03500;

XX 26-SEP-2001 (first entry)

DT Human TWEAKR receptor-Fc (TWEAKR-Fc) fusion polypeptide.

DE TWEAKR receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;
KW corneal graft neovascularisation; psoriasis; metastatic condition;
KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
KW preneoplastic condition; myocardial angiogenesis; wound granulation;
KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;
KW peripheral atherosclerosis; human IgG1; TWEAKR-Fc; fusion protein.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers
FT Peptide 1..27
/note= "Signal peptide"

FT Protein 28..309
/note= "Mature human TWEAKR-Fc fusion protein."

FT Domain 28..79
/note= "Specifically referred to in Claim 11"

FT Region 80..81
/note= "From TWEAKR extracellular domain. Specifically referred to in Claims 4 and 10"

FT Region /note= "From a BglII cloning site"
FT 82...309
FT /note= "Fc portion"
XX
PN W0200145730-A2.
XX
XX 28-JUN-2001.
XX
XX 19-DEC-2000; 2000WO-US34755.
XX
XX 20-DEC-1999; 99US-0172878.
PR 10-MAY-2000; 2000US-0203347.
XX
XX (IMMV) IMMUNEX CORP.
XX
XX Wiley SR;
PI
XX WPI; 2001-417975/44.
DR N-PSDB; AAS03965.
XX
XX Modulating angiogenesis in a mammal for treating diseases mediated by
PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
PT peripheral tissue, by administering antagonist or agonist of TWEAK
PT receptor
XX
XX Claim 4; Page 45-46; 46pp; English.
XX
XX The sequence represents a fusion protein consisting of the human TWEAK
CC receptor (TWEAKR) protein extracellular domain fused to an Fc portion
CC from human IgG1. This fusion protein, TWEAKR-Fc, is used in the
CC preparation of TWEAKR agonists and antagonists. The TWEAK protein is a
CC member of the tumour necrosis factor (TNF) family and induces
CC angiogenesis. TWEAKR may therefore be used to screen for and develop
CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
CC used in the treatment and diagnosis of human disease. The disorders
CC mediated by angiogenesis include ocular disorders characterised by ocular
CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,
CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
CC rubeosis, uveitis, macular degeneration and corneal graft
CC neovascularisation, and inflammatory diseases such as arthritis,
CC rheumatism and psoriasis. Other treatable diseases include malignant and
CC metastatic conditions such as sarcomas and carcinomas, benign tumours and
CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
CC scleroderma, vascular adhesions, atherosclerotic plaque
CC neovascularisation, telangiectasia, wound granulation, coronary
CC atherosclerosis, peripheral atherosclerosis and ischaemia.
XX
SQ Sequence 309 AA;

Query Match 100.0%; Score 1551; DB 22; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-111;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EQAPGTAPCRSGSSWADLCKMDCASCARPHSDFCGLGCAAAPPPAPFLLWRSCKDTHT 60
Db 28 eqapgtapcrgsswsadlckmdcascarpshdflcgaapppapfllwrsckdtht 87
QY 61 CPPCPAPEAGAPSVFLFPKPKDLMISRTPEVTVVVDVSHDEPVEKFNWYDGVVEH 120
Db 88 cpcpapeagapsvflfpkpkdmlisrtpevtcvvvdvshdpevkfnwydgvveh 147
QY 121 NAKTKPREQYNSTRYRVSVTLVHLQDWLNGKEYCKVSKNALPAPTEKTSKAKGPRE 180
Db 148 naktkpreeqynstryrvsvtlvhlqdwlngkeyckvsknalpapektskagpre 207
QY 181 PQVYTLPPSREEMTKNOVSLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPPVLDSGCSFF 240
Db 208 pqvylppsreemtknqysitclvkgfypsdiavewesngopennykttppvltdsgsff 267
QY 241 LYSKLTVDKRWQGNVFCSSVMHEALHNYHTOKSLSPGK 282
Db 268 lyskltvdksrwqgnvfscsvmhealhnhytqkslslspgk 309

RESULT 2

AAV43139
ID AAY43139 standard; Protein; 451 AA.
XX
XX AAY43139;
XX
XX 23-DEC-1999 (first entry)
XX
XX NAIL-Fc protein sequence.
XX
XX NAIL; human; natural killer cell; NK Cell Activation Inducing Ligand;
KW immunogen; B cell stimulation; cancer cell proliferation; inhibitor;
KW CD48; dendritic cell; cytotoxic T cell; cancer; autoimmune disease;
KW therapy; fusion protein.
XX
XX Homo sapiens.
OS
OS Synthetic.
XX
XX WO9950297-A1.
XX
XX 07-OCT-1999.
XX
XX 23-MAR-1999; 99WO-US06215.
XX
XX 27-MAR-1998; 98US-0079845.
PR 17-AUG-1998; 98US-0096750.
XX
XX (IMMV) IMMUNEX CORP.
XX
XX Kubin MZ, Goodwin RG;
XX
XX WPI; 1999-591275/50.
XX
XX New NK cell activation ligand (NAIL) polypeptides stimulate B, NK and T
PT cells and eliminate cancer cells
XX
XX Disclosure; Page 25-26; 113pp; English.
XX
XX This sequence is a fusion protein containing the human NK (Natural
CC Killer) Cell Activation Inducing Ligand (NAIL) protein of the invention.
CC An immunogenic composition containing a NAIL polypeptide may be
CC useful for: stimulating B cells; stimulating NK cells; inhibiting the
CC proliferation of cancer cells; chelating soluble CD48 in a patient;
CC inhibiting the binding of CD48 with NAIL on the cell surface in a
CC patient; inhibiting the binding of NAIL with CD48 on the cell surface in
CC a patient; stimulating dendritic cells in a patient; stimulating NK cells
CC in a patient; stimulating cytotoxic T cells in a patient; inhibiting the
CC stimulation of NK cells or cytotoxic T cells in a patient; inhibiting the
CC stimulation of B cells or dendritic cells in a patient. The nucleic acids
CC of the invention can be used as probes to identify nucleic acid encoding
CC proteins having NAIL activity and to inhibit expression of polypeptides
CC encoded by the NAIL gene. Another use of the polypeptide is as a research
CC tool for studying the biological effects that result from inhibiting
CC NAIL/CD48 interactions on different cell types and to study cell signal
CC transduction. Antibodies against NAIL can be used to detect the presence
CC of NAIL in a sample and can be used therapeutically to bind to NAIL and
CC inhibit its activity in vivo. The antibodies can also be useful in the
CC diagnosis of pathological states that result in overexpression or
CC underexpression of NAIL such as in cancers and autoimmune diseases. The
CC invention provides polypeptides and antibodies suitable for use in
CC studies of the modulation of NK, T cell and B cell activity and in the
CC selection of specific cell types. Furthermore, it provides polypeptides
CC and antibodies suitable for use in studies of CD48, where a lack of
CC suitable reagents have previously been unavailable.
XX
XX Sequence 451 AA;

Query Match 81.1%; Score 1258; DB 20; Length 451;
Best Local Similarity 86.3%; Pred. No. 8.8e-89;
Matches 240; Conservative 3; Mismatches 15; Indels 20; Gaps 3;

```
QY 6 TAPCSRGSSW-SADLDKCMDCASCRRPHSDFLGCAAPAPAPFRLMWSCKTHTCCPC 64
Db 193 tcnvsnpsweshntlnltqdcgna----hgef-----rrscdkthtccpc 233
QY 65 PAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKT 124
Db 234 papeaegapsvflfpkpkdtlmisrtpevtcvvdvshedpevkfnwydgvvhnakt 293
QY 125 KPREGQYNSTYRVSVLTVLHQDWLNGKEYCKVKSNKALPAPIETKISKAKQPREPVY 184
Db 294 kpreeqynstyrvsvltvltlqhgdwlngkeyckvksnkalpapietkiskakgprepv 353
QY 185 TLPPSREEMTKNOVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLY 244
Db 354 tlppreemtknqvsltcclvkgfypsdiavewesngqpennnykttppvldsdgsfl 413
QY 245 LVVDKSRMQQGNVFCSCVMHEALHNHYTQKSLSLSPGK 282
Db 414 ltvdksrwgqgnvfscsvmhealhnhytqkslsispkg 451

RESULT 3
AAR81882
ID AAR81882 standard; Protein: 438 AA.
XX
AC AAR81882;
XX
DT 30-MAR-1996 (first entry)
XX
DE Plasmid pBC406/OX40/Fc* encoding an OX40/Fc mutein protein.
XX
KW OX40; OX40-L; cytokine; cell surface molecule; plasmid;
KW pDC406/OX40/Fc*; membrane glycoprotein.
XX
OS Synthetic.
XX
PN US5457035-A.
XX
PD 10-OCT-1995.
XX
PF 23-JUL-1993; 93US-0097827.
XX
PR 23-JUL-1993; 93US-0097827.
XX
PA (IMMV ) IMMUNEX CORP.
XX
PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
XX
DR WPI: 1995-357992/46.
DR N-PSDB; AAT00829.
XX
XX New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT and host cells, used to produce recombinant ligand used in e.g.
PT prim. T cell culture, to modulate immune response etc.
XX
PS Example 2; Column 35-38; 26pp; English.
XX
CC This plasmid encodes an OX40/Fc antibody fragment mutein protein,
CC and is used to express a soluble OX40/Fc mutein fusion protein for
CC use in detecting cDNA clones encoding a OX40 ligand. The Fc
CC fragment may be derived from human IgG1, and the plasmid may be
CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
CC line. Culture supernatant was purified by affinity chromatography
CC and this was used, together with labeled goat anti-human IgG to
CC screen various cell lines.
XX
SQ Sequence 438 AA;
```

Query Match 80.9%; Score 1255; DB 16; Length 438;
Best Local Similarity 65.8%; Pred. No. 1.4e-88;
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;

```
QY 8 PCSRG-SSWSADLDKCMDCAS-----CRARP-----HSDFCLG- 39
Db 62 pceetfyneavnydtckgctqcnhrsgselkqncptqdtvcrpqtgprqdgsgyklgv 121
QY 40 -CAAAPAPPAPF-----RLLW----- 52
Db 122 devpcppghfsgnngackpwtncnctlsqkqtrhpasdsldavcedrslatlilwetqrpt 181
QY 53 -----RSCDKTHTCPPCPAPEAGAPSVFLFPKPKDT 85
Db 182 frptvtqvstvtwprtselpstptllveprscdkthtccpcpapeaegapsvflfpkpkdt 241
QY 86 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTTPREOYNSTYRVSVLTVLH 145
Db 242 lmisrtpevtcvvdvshedpevkfnwydgvvhnaktkpreeqynstyrvsvltvln 301
QY 146 QDWLNGKEYCKVKSNKALPAPIETKISKAKQPREPVYTLPPSREEMTKNOVSLTCLVK 205
Db 302 qdwlngkeyckvksnkalpapietkiskakgprepvyltppsrdelcknqvsclclvk 361
QY 206 GFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSLKLTVDKSRWQGNVFCSCVMHE 265
Db 362 gfypsdiavewesngqpennnykttppvldsdgsfflyskltvdksrwgqgnvfscsvmhe 421
QY 266 ALHNHYTQKSLSLSPGK 282
Db 422 alhnhytqkslsispkg 438

RESULT 4
AAW48976
ID AAW48976 standard; Protein: 438 AA.
XX
AC AAW48976;
XX
DT 25-SEP-1998 (first entry)
XX
DE OX40/Fc mutein.
XX
KW OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;
KW chimeric.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Mus sp.
XX
FH Key Location/Qualifiers
FT Region 1..206
FT /note= "Extracellular domain of mouse OX40"
FT Region 207..438
FT /note= "Mutant Fc region of human IgG1 antibody"
FT Misc-difference 225
FT /note= "changed from Leu in wild-type to Ala in
FT mutant"
FT Misc-difference 226
FT /note= "changed from Leu in wild-type to Gly in
FT mutant"
FT Misc-difference 228
FT /note= "changed from Gly in wild-type to Ala in
FT mutant"
XX
US5783665-A.
XX
PD 21-JUL-1998.
XX
PF 22-JUN-1995; 95US-0494574.
XX
PR 23-JUL-1993; 93US-0097827.
PR 22-JUN-1995; 95US-0494574.
XX
PA (IMMV ) IMMUNEX CORP.
XX
PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
```

```
XX WPI; 1998-427099/36.
DR N-PSDB; AAV32636.
XX
XX Purified polypeptide OX-40 ligands - for co-stimulation of T-cell
PT production and binding assays for OX-40 and homologues
XX
XX Example 2; Col 37-40; 26pp; English.
XX
XX The present sequence represents the OX40/Fc fusion protein that
CC contains the extracellular domain of mouse OX40 fused to the mutated
CC FC region of the human IgG1 antibody. The fusion protein was used
CC for detecting cDNA clones encoding an OX40 ligand. The invention
CC claims for a murine OX40-L cytokine (AAW48975) that binds to the murine
CC T cell antigen, OX40. The OX40-L protein is claimed to be useful for
CC co-stimulation of T-cell production and in binding assays for
CC detecting OX40 or its homologues. The OX40-L protein is also claimed
CC to generate a TH-2 immune response.
XX
XX Sequence 438 AA;
XX
XX Query Match 80.9%; Score 1255; DB 19; length 438;
XX Best Local Similarity 65.8%; Pred. NO. 1.4e-88;
XX Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;
XX
QY 8 PCSRG-SSWSADLDKCMDCAS-----CRARP-----HSDFCLG- 39
DB 62 pctgfyneavnydtkqctqcnhrsgselkqncptqdtvcrprgtqrgsgykglv 121
QY 40 -CAAAPAPAP-----RSCDKTHTCCPCPAPAEAGAPSVFLFPKPKDT 52
DB 122 dcvcpgphspgnnqackpwtntclsgkqtrhpasdsldavcedslatlilwetqrpt 181
QY 53 -----RSCDKTHTCCPCPAPAEAGAPSVFLFPKPKDT 85
DB 182 frptvgsttwprtselfptptlveprsdckthtccpcpapeagapsvflfppkpkdt 241
QY 86 LMTSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVHL 145
DB 242 lmsrtpcvvvdvshedpevkfnwvvgvvevhnaktprceeqynstyrvvsvltvln 301
QY 146 QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVK 205
DB 302 qdwlngkeykckvsnkalpapiektskagkqpreqvyltppsrdektknqysltclvk 361
QY 206 GFVPSDIAVWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHE 265
DB 362 gfypsdiavewesngqpennnykttppvldsdgsfflyskltvdksrwqggnvfscvmhe 421
QY 266 ALHNHYTQKSLSLSPGK 282
DB 422 alnhytqkslsispgk 438
XX
RESULT 5
AAV13463
ID AAY13463 standard; Protein; 660 AA.
XX
XX AAY13463;
XX
XX 26-JUL-1999 (first entry)
XX
XX Ectromelia A39R semaphorin polypeptide.
XX
XX VESPR; viral-encoded semaphorin protein receptor; semaphorin; IL-12;
XX inflammatory disease; viral infection; immune regulation; interleukin-12;
XX Th1 type immune response; cancer; autoimmune disease; A39.
XX
XX Ectromelia virus.
XX
XX WO921997-A1.
XX
```

```
PD 06-MAY-1999.
XX
XX 28-OCT-1998; 98WO-US22879.
XX
XX 28-OCT-1997; 97US-0112009.
XX
XX (IMMV ) IMMUNEX CORP.
XX
XX Comeau MR, Dubose RF, Johnson RS, Spriggs MK;
XX
XX WPI; 1999-326704/27.
XX N-PSDB; AAX55565.
XX
XX Viral-encoded protein receptors useful for treating inflammatory
XX diseases
XX
XX Example 1; Page 63-66; 73pp; English.
XX
XX The invention relates to a VESPR (viral-encoded semaphorin protein
XX receptor) polypeptide that can bind semaphorins. The VESPR polypeptides
XX are used to treat inflammatory diseases; when immobilized, to purify and
XX separate semaphorins or cells that express them. They are useful as
XX reagents for detecting, or measuring biological activity of, optionally
XX modified semaphorins; as carriers for delivering diagnostic and
XX therapeutic agents to semaphorin-positive cells; for studying the role of
XX VESPR in viral infection and immune regulation. When administered, in
XX combination with vaccination, together with a semaphorin, VESPR may
XX stimulate production of interleukin-12 (IL-12) which acts as an adjuvant
XX to induce a more persistent cellular immune response (of Th1 type),
XX including a curative response against aggressive, micrometastatic
XX cancers. Antibodies against VESPR can be used to treat autoimmune
XX diseases where an inflammatory response follows presentation of self-
XX antigens by T cells. Sense and antisense fragments of nucleic acid
XX encoding VESPR can be used to block expression of the polypeptide. The
XX present sequence represents an eromelia A39 semaphorin protein. The
XX native VESPR polypeptide is isolated from the human cells expressing the
XX receptor by using an Ectromelia virus A39 semaphorin/fc Fusion protein.
XX
XX Sequence 660 AA;
XX
XX Query Match 80.6%; Score 1250; DB 20; Length 660;
XX Best Local Similarity 100.0%; Pred. NO. 5.8e-88;
XX Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 53 RSCDKTHTCCPCPAPAEAGAPSVFLFPKPKDTLMTSRTPEVTCVVVDVSHEDPEVKFNW 112
DB 34 rscdkthtccpcpapeagapsvflfppkpkdtlmsrtpevtcvvvdvshedpevkfnw 93
QY 113 YDGVGVHNAKTKPREEQYNSTYRVVSVLTVHLQDWLNGKEYKCKVSNKALPAPIEKTIS 172
DB 94 ydgvgevhnaktprceeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiektis 153
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPV 232
DB 154 kakgqpreqvyltppsrreemtknqvsltcclvkgyfypsdiavewesngqpennnykttppv 213
QY 233 LDSGSGFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHYTQKSLSLSPGK 282
DB 214 ldsdgsfflyskltvdksrwqggnvfscsvmhealnhytqkslsispgk 263
XX
RESULT 6
AAB28523
ID AAB28523 standard; Protein; 660 AA.
XX
XX AAB28523;
XX
XX 07-FEB-2001 (first entry)
XX
XX Ectromelia A39R semaphorin.
XX
XX Ectromelia; semaphorin; VESPR; viral encoded semaphorin protein receptor;
```


KW anti-inflammatory; semaphorin modulator; rheumatoid arthritis;
KW inflammation; immune regulation; viral infection.
XX Ectromelia virus.
PN US6130068-A.
XX 10-OCT-2000.
XX 28-OCT-1998; 98US-0181706.
XX 26-OCT-1998; 98US-0112009.
XX (IMMV) IMMUNEX CORP.
PI Comeau MR, Johnson RS, Spriggs MK, Dubose RF;
XX WPI; 2000-646753/62.
DR N-PSDB; AAC63726.
XX New Viral Encoded Semaphorin Protein Receptor DNA and polypeptides,
PT useful for treating inflammation or inflammatory diseases, e.g.
PT rheumatoid arthritis -
XX Example 1; Column 55-60; 32pp; English.
XX The present sequence is provided in an invention relating to novel
CC semaphorin receptor polypeptides designated Viral Encoded Semaphorin
CC Protein Receptor (VESPR). The DNA encoding the VESPR is useful for
CC constructing expression vectors that produce the VESPR polypeptide. The
CC VESPR DNA and polypeptide are useful for treating inflammation and
CC inflammatory diseases, e.g. rheumatoid arthritis, a disease associated
CC with chronic inflammation of the synovial tissue. The VESPR polypeptide
CC is useful for measuring the biological activity of semaphorin proteins
CC in terms of their binding affinity for VESPR. It is also useful as a
CC research tool for studying the role that the receptor, in conjunction
CC with semaphorins, may play in immune regulation and viral infection. The
CC VESPR polypeptides are also useful as reagents that may be employed by
CC those conducting quality assurance studies, e.g. to monitor shelf life
CC and stability of semaphorin protein under different conditions.
XX
SQ Sequence 660 AA;

Query Match 80.6%; Score 1250; DB 21; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.8e-88;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112
DB 34 rscdkthtcppcpapeagapsvflfpkpkdtlmisrtpevtcvvdvshedpevkfnw 93
QY 113 YVDGVEVHNATKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172
DB 94 yvdgvevhnatkpreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiektis 153
QY 173 KAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPV 232
DB 154 kagprepqvylppstreemtknqvsitclvkgfypsdiavewesngopennykttppv 213
QY 233 LDSGCSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK 282
DB 214 ldsdgsfflyskltvdksrwqqgnvfscsvmhealhnhytqkslsispkg 263

RESULT 7
AAB70132
ID AAB70132 standard; Protein; 660 AA.
XX
AC AAB70132;
XX
DT 21-MAY-2001 (first entry)
XX

DE Ectromelia A39R semaphorin.
XX
KW Vaccinia virus; Ectromelia; A39R; semaphorin; VESPR;
KW viral encoded semaphorin protein receptor; antiinflammatory;
KW inflammation.
XX
OS Vaccinia virus.
PN US6187909-B1.
XX 13-FEB-2001.
PD 10-DEC-1999; 99US-0459066.
PF 26-OCT-1998; 98US-0112009.
PR 28-OCT-1998; 98US-0181706.
PR 28-OCT-1997; 97US-0958598.
XX (IMMV) IMMUNEX CORP.
XX Spriggs MK, Comeau MR, Dubose RF, Johnson RS;
XX WPI; 2001-256141/26.
DR N-PSDB; AAF76950.
XX
XX New viral encoded semaphorin protein receptor polypeptides useful for
PT treating inflammation or diseases associated with pro-inflammatory
PT activity of a semaphorin ligand -
PS Example 1; Column 55-60; 32pp; English.
XX The DNA encoding the present sequence was used to prepare an Ectromelia
CC Semaphorin/Fc fusion protein. The fusion protein was used in the
CC isolation of a human viral encoded semaphorin protein receptor (VESPR).
CC The VESPR polypeptides are useful for treating inflammation, or diseases
CC associated with pro-inflammatory activity of a semaphorin ligand. The
CC polypeptides are also useful as carriers for delivering diagnostic or
CC therapeutic agents to cells expressing semaphorins.
XX
SQ Sequence 660 AA;

Query Match 80.6%; Score 1250; DB 22; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.8e-88;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112
DB 34 rscdkthtcppcpapeagapsvflfpkpkdtlmisrtpevtcvvdvshedpevkfnw 93
QY 113 YVDGVEVHNATKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 172
DB 94 yvdgvevhnatkpreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiektis 153
QY 173 KAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPV 232
DB 154 kagprepqvylppstreemtknqvsitclvkgfypsdiavewesngopennykttppv 213
QY 233 LDSGCSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK 282
DB 214 ldsdgsfflyskltvdksrwqqgnvfscsvmhealhnhytqkslsispkg 263

RESULT 8
AAB31694
ID AAB31694 standard; Protein; 660 AA.
XX
AC AAB31694;
XX
DT 30-APR-2001 (first entry)
XX
DE Amino acid sequence of the Ectromelia virus A39R semaphorin.
XX

KW Viral encoded semaphorin protein receptor; VESPR; semaphorin; interferon;
KW viral semaphorin ligand; Ectromelia A39R; Staphylococcus aureus; IL-12;
KW interleukin-12; dendritic cell; cytokine production; immunomodulator;
KW proinflammatory cytokine; Th1 cell differentiation; inflammation;
KW inflammatory disease; tumour.

XX Ectromelia virus.
XX US6174689-B1.
XX 16-JAN-2001. 9905-0458791.
XX 10-DEC-1999; 9905-0458791.
XX 26-OCT-1998; 9805-0112009.
XX 28-OCT-1998; 9805-0181706.
XX 27-OCT-1997; 9705-0958598.
XX (IMMV) IMMUNEX CORP.

PI Spriggs MK, Comeau MR, Dubose RF, Johnson RS;
XX WPI; 2001-167789/17.
XX N-PSDB; AAF25254.
XX Screening for binding to viral encoded semaphorin protein receptor,
XX comprises contacting a mixture containing semaphorin, or cells that
XX express semaphorin with protein, and detecting binding to protein -
XX Example 1; Column 55-60; 32pp; English.

XX The present sequenc represents an Ectromelia virus A39R semaphorin
XX ligand. The protein was used to isolate a viral encoded semaphorin
XX protein receptor (VESPR) polypeptide from human cells. Semaphorins
XX interact with their membrane bound receptors to synergise with interferon
XX and Staphylococcus aureus (type C) in the production of interleukin-12
XX (IL-12) from dendritic cells. VESPR can therefore be used to induce
XX IL-12 production which in turn promotes cytokine production. IL-12 is
XX a proinflammatory cytokine and an immunomodulator. A soluble VESPR can
XX be used to antagonise IL-12 and downregulate Th1 cell differentiation.
XX VESPR is therefore useful for treating inflammation and inflammatory
XX diseases. VESPR ligands may be used to induce an immune response against
XX aggressive tumours.

XX Sequence 660 AA;

Query Match 80.6%; Score 1250; DB 22; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.8e-88;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 RSCDKTHCPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVFNW 112
DB 34 rscdkthtccpapeagapsvflppkpkdtlmisrtpevtcvvvdvshedpevkfnw 93
QY 113 YVDGVEVHNKTKPREQYNTSTYRVSVLTVLHODWLNKREYCKVSNKALPAPIKTIS 172
DB 94 yvdgvevhnaktkpreeqyntstyrvsvlvtvhgdwlngkeyckvsnkalpapiektis 153
QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 232
DB 154 kagqprepqvtytlppsreemtqnqvscltclvkgyfypsdiavewesngqpennnykttppv 213
QY 233 LDSGSRFLYSKLTVDKSRWQQGNVFCSVNHEALHNHYTQKSLSLSPGK 282
DB 214 ldsdgsrflyskltvdksrwqgnvfscsvnmhealhnhytqkslsislspgk 263

RESULT 9
ID AAY15239 standard; Protein; 450 AA.
XX AAY15239;
AC "

XX 04-NOV-1999 (first entry)
DT ULBP1-Fc fusion polypeptide amino acid sequence.
XX cell surface glycoprotein; lymphoma; assay; retinitis pigmentosa;
XX diabetes mellitus; progressive pseudo rheumatoid arthropathy;
XX muscular dystrophy; congenital merosin-deficiency; cardiomyopathy.
XX fusion polypeptide.

XX Homo sapiens.
XX Synthetic.
XX WO9931241-A1.
XX 24-JUN-1999.
XX 17-DEC-1998; 98WO-US27048.
XX 15-JUL-1998; 98US-0092946.
XX 17-DEC-1997; 97US-0069857.
XX (IMMV) IMMUNEX CORP.

XX Cosman DJ, Fanslow WC, Mullberg JH;
XX WPI; 1999-493777/41.
XX DNA encoding human B cell lymphocyte cell surface glycoproteins
XX Disclosure; Page 29; 104pp; English.

XX This is the amino acid sequence of the ULBP1-Fc fusion polypeptide. The
XX ULBP-1 and ULBP-2 polypeptides can be expressed as Fc fusion proteins
XX using an Fc mucrein to provide fused polypeptides, such as ULBP1-Fc.
XX The polypeptides ULBP-1 and -2 bind UL16-Fc. In addition, they bind to a
XX number of human cell types, including mitogen-stimulated human T cells
XX and natural killer (NK) cells. ULBP-Fc proteins bind to K299 cells, an
XX anaplastic lymphoma. The ULBP proteins can therefore be used as markers
XX to detect cancer, to enhance IFN- gamma production, NK cell proliferation
XX and CTL activity, to purify proteins and measure their activity. The
XX polypeptides and their fragments can also be used as delivery and
XX therapeutic agents, for rational drug design, as research reagents,
XX controls for peptide fragmentation, molecular weight/isoelectric focusing
XX markers, identification of unknown proteins and also for preparation of
XX antibodies.
XX The antibodies can be used in assays to detect the presence of ULBP
XX proteins, in vitro or in vivo, as well as for use in purification of ULBP
XX proteins. The ULBP DNA sequences can be used as probes to identify
XX homologues, to identify human chromosome number 6, to map and identify
XX genes, especially associated with certain diseases, syndromes or other
XX conditions on chromosome 6, as single-stranded sense or antisense
XX oligonucleotides, to inhibit expression of ULBP polypeptides, to help
XX detect defective genes in an individual and for gene therapy.
XX Diseases, syndromes and conditions associated with human chromosome
XX 6 include Retinitis pigmentosa (6q14-q21), Diabetes mellitus
XX (insulin-dependent) (6q21), progressive pseudorheumatoid arthropathy of
XX childhood (6q22), muscular dystrophy (congenital merosin-deficient)
XX (6q22-q23), and cardiomyopathy (dilated) (6q23).

XX Sequence 450 AA;

Query Match 80.4%; Score 1247; DB 20; Length 450;
Best Local Similarity 96.3%; Pred. No. 6.2e-88;
Matches 232; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 44 PP--APFRLWRSCDKTHCPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVDV 101
DB 210 ppslapgttprscdkthtccpapeagapsvflppkpkdtlmisrtpevtcvvvdv 269
QY 102 SHEDPEYKFNWYVDGVEVHNKTKPREEQYNTSTYRVSVLTVLHODWLNKREYCKVSNK 161
|||||

Db 270 shedpevfwnvydvgevhnaaktppreeqynstyrsvsvltvlhgdwlngkeyckvsnk 329
QY 162 ALPAPEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVESHNGQ 221
Db 330 alpapietiskakgqprepqvvtlppsrdeitknqvsitclvkvfypsdiaveshngq 389
QY 222 PENNYKTTTPPVLDSGSEFELYSLTKLVKDSRWQGNVFCSCVMHEALHNHYTKQSLSPG 281
Db 390 pennykttppvlidsgsflyskltvdkvskrwqgnvfscsvmhealhnhytqkslspsg 449
QY 282 K 282
Db 450 k 450

RESULT 10
AAV15240
ID AAV15240 standard; Protein; 453 AA.
XX
AC AAV15240;
XX
DT 04-NOV-1999 (first entry)
XX
DE ULBP2-Fc fusion polypeptide amino acid sequence.
XX
KW cell surface glycoprotein; lymphoma; assay; retinitis pigmentosa;
KW diabetes mellitus; progressive pseudo rheumatoid arthropathy;
KW muscular dystrophy; congenital merosin-deficiency; cardiomyopathy.
KW fusion polypeptide.
XX
OS Homo sapiens.
XX
PN WO9931241-A1.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-US27048.
XX
PR 15-JUL-1998; 98US-0092946.
PR 17-DEC-1997; 97US-0069857.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Cosman DJ, Fanslow WC, Mullberg JH;
XX
DR WPI; 1999-493777/41.
XX
PT DNA encoding human B cell lymphocyte cell surface glycoproteins
XX
PS Disclosure; Page 29; 104pp; English.
XX
CC This is the amino acid sequence of the ULBP2-Fc fusion polypeptide. The
CC ULBP-1 and ULBP-2 polypeptides can be expressed as Fc fusion proteins
CC using an Fc mutain to provide fused polypeptides, such as ULBP2-Fc.
CC The polypeptides ULBP-1 and -2 bind UL16-Fc. In addition, they bind to a
CC number of human cell types, including mitogen-stimulated human T cells
CC and natural killer (NK) cells. ULBP-Fc proteins bind to K299 cells, an
CC anaplastic lymphoma. The ULBP proteins can therefore be used as markers
CC to detect cancer, to enhance IFN- gamma production, NK cell proliferation
CC and CTL activity, to purify proteins and measure their activity. The
CC polypeptides and their fragments can also be used as delivery and
CC therapeutic agents, for rational drug design, as research reagents,
CC controls for peptide fragmentation, molecular weight/isoelectric focusing
CC markers, identification of unknown proteins and also for preparation of
CC antibodies.
CC The antibodies can be used in assays to detect the presence of ULBP
CC proteins, in vitro or in vivo, as well as for use in purification of ULBP
CC proteins. The ULBP DNA sequences can be used as probes to identify
CC homologues, to identify human chromosome number 6, to map and identify
CC genes, especially associated with certain diseases, syndromes or other
CC conditions on chromosome 6, as single-stranded sense or antisense
CC oligonucleotides, to inhibit expression of ULBP polypeptides, to help
CC detect defective genes in an individual and for gene therapy.

CC Diseases, syndromes and conditions associated with human chromosome
CC 6 include Retinitis pigmentosa (6q14-q21), Diabetes mellitus
CC (insulin-dependent) (6q21), progressive pseudorheumatoid arthropathy of
CC childhood (6q22), Muscular dystrophy (congenital merosin-deficient)
CC (6q22-q23), and cardiomyopathy (dilated) (6q23).
XX
SQ Sequence 453 AA;
Query Match 80.4%; Score 1247; DB 20; Length 453;
Best Local Similarity 91.5%; Pred. No. 6.2e-88;
Matches 236; Conservative 3; Mismatches 9; Indels 10; Gaps 2;
QY 35 DFCLG-----CAAAPPA---PFRLLWRSCDKTHTCPCPAPEAGAPSVFLFPKPKD 84
Db 196 dfllmgmdstlepsagaplamssgttqlrrscdkttccpapeagapsvflfpkpkd 255
QY 85 TLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNATKPREQYNSTYRVVSVLTVTL 144
Db 256 tlmisrtpevtcvvdvshedpevkfnwydvgevhnaaktppreeqynstyrsvsvltvl 315
QY 145 HQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLV 204
Db 316 hgdwlngkeyckvsnkalpapietiskakgqprepqvvtlppsrdeitknqvsitclv 375
QY 205 KGFYPSDIAVESHNGQPENNYKTTTPPVLDSGSEFELYSLTKLVKDSRWQGNVFCSCVMH 264
Db 376 kgfypsdiaveshngqpennnykttppvlidsgsflyskltvdkvskrwqgnvfscsvmh 435
QY 265 EALHNHYTKQSLSPGK 282
Db 436 ealhnhytqkslspsgk 453

RESULT 11
AAW14764
ID AAW14764 standard; Protein; 424 AA.
XX
AC AAW14764;
XX
DT 11-JUN-1997 (first entry)
XX
DE Human soluble kit ligand-IgG fusion protein.
XX
KW Kit ligand; c-kit proto-oncogene; cytokine; growth factor;
KW haematopoietic cell; cell proliferation; stem cell; anaemia;
KW thrombocytopaenia; therapy; IgG1.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..25 /label= Sig_peptide
FT /note= "KL signal peptide"
FT Protein 26..424 /label= Mat_protein
FT /label= Mat_protein
FT /note= "human KL-IgG fusion"
XX
PN WO9526199-A1.
XX
PD 05-OCT-1995.
XX
PF 28-MAR-1995; 95WO-US03866.
XX
PR 28-MAR-1994; 94US-0220379.
XX (CYTO-) CYTOMED INC.
XX Lobell RB, Nocka KH;
XX WPI; 1995-351198/45.
XX N-PSDB; AAT63109.
XX

PT Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit
 PT increased activity in promoting cell proliferation
 XX
 XX
 PS Claim 10; Page 43-44; 88pp; English.

XX A fusion protein (AAW14764) between human soluble kit ligand (KL)
 CC (see also AAW14761) and a human IgG1 heavy chain can be transiently
 CC expressed in COS cells transfected with a human KL-Ig cDNA
 CC construct (AAT63109) in vector CDM8; a corrected KL-Ig construct
 CC (AAW14765) has also been prepd. KL-Ig can also be produced as a
 CC dimer stabilised by intermolecular disulphide bonds or a peptide
 CC linker. The stabilised KL-Ig dimers have a more favorable cell
 CC proliferation: mast cell activation ratio than native KL and can
 CC stimulate haematopoietic recovery or stem cell/progenitor cell
 CC mobilisation with less toxicity.

XX Sequence 424 AA;

Query Match 80.1%; Score 1243; DB 16; Length 424;
 Best Local Similarity 84.6%; Pred. No. 1.2e-87;
 Matches 237; Conservative 6; Mismatches 25; Indels 12; Gaps 2;

QY 3 APGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLCAAAPAPFRLLRSCDKTHTCP 62
 Db 157 asetsdcvssstlspekds-----rvsvtkpflmpvvaadpep-----kscdkthtcp 204
 QY 63 PCPAPEAEGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 122
 Db 205 pcpapellggpsvflfpkpkdtlmisrtpevtcvvvdshedpevkfnwvvgvghna 264
 QY 123 KTRPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIETKISKAKGQPREPQ 182
 Db 265 ktkpreeqynstyrvvsvltvllhqdwlngkeyckvknsnkalspapiektiskakgqprepq 324
 QY 183 VYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 242
 Db 325 vytlppsrdeitknqvsaltclvkgfypsdiavewesngqpennnykttppvldsdgsfely 384
 QY 243 SKLTVDSRWQOGNVFSCSVHMEALHNHYTQKSLSLSPGK 282
 Db 385 skltvdksrwqggnvfscsvmhleahhnytkqkslsispgk 424

RESULT 12

AAW14765
 ID AAW14765 standard; Protein; 424 AA.

XX AAW14765;

XX 11-JUN-1997 (first entry)

XX Human soluble kit ligand-IgG fusion protein (corrected).

XX Kit ligand; c-kit proto-oncogene; cytokine; growth factor;
 KW haematopoietic cell; cell proliferation; stem cell; anaemia;
 KW thrombocytopaenia; therapy; IgG1.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 1..25
 FT /label= Sig_peptide
 FT /note= "KL signal peptide"
 FT Protein 26..424
 FT /label= Mat_protein
 FT /note= "human KL-Ig fusion"

PN W09526199-A1.

XX 05-OCT-1995.

XX 28-MAR-1995; 95WO-US03866.

XX 28-MAR-1994; 94US-0220379.
 XX (CYTO-) CYTOMED INC.
 XX
 XX Lobell RB, Nocka KH;
 XX WPI: 1995-351198/45.
 DR N-PSDB; AAT63110.

XX Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit
 PT increased activity in promoting cell proliferation
 XX
 XX Claim 10; Page 46-48; 88pp; English.

XX A fusion protein (AAW14765) between human soluble kit ligand (KL)
 CC (see also AAW14761) and a human IgG1 heavy chain can be transiently
 CC expressed in COS cells transfected with a human KL-Ig cDNA
 CC construct (AAT63110) in vector CDM8. KL-Ig can also be produced as a
 CC dimer stabilised by intermolecular disulphide bonds or a peptide
 CC linker. The stabilised KL-Ig dimers have a more favorable cell
 CC proliferation: mast cell activation ratio than native KL and can
 CC stimulate haematopoietic recovery or stem cell/progenitor cell
 CC mobilisation with less toxicity.

XX Sequence 424 AA;

Query Match 80.1%; Score 1243; DB 16; Length 424;
 Best Local Similarity 84.6%; Pred. No. 1.2e-87;
 Matches 237; Conservative 6; Mismatches 25; Indels 12; Gaps 2;

QY 3 APGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLCAAAPAPFRLLRSCDKTHTCP 62
 Db 157 asetsdcvssstlspekds-----rvsvtkpflmpvvaadpep-----kscdkthtcp 204
 QY 63 PCPAPEAEGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 122
 Db 205 pcpapellggpsvflfpkpkdtlmisrtpevtcvvvdshedpevkfnwvvgvghna 264
 QY 123 KTRPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIETKISKAKGQPREPQ 182
 Db 265 ktkpreeqynstyrvvsvltvllhqdwlngkeyckvknsnkalspapiektiskakgqprepq 324
 QY 183 VYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 242
 Db 325 vytlppsrdeitknqvsaltclvkgfypsdiavewesngqpennnykttppvldsdgsfely 384
 QY 243 SKLTVDSRWQOGNVFSCSVHMEALHNHYTQKSLSLSPGK 282
 Db 385 skltvdksrwqggnvfscsvmhleahhnytkqkslsispgk 424

RESULT 13

AAW60037

ID AAW60037 standard; Protein; 376 AA.

XX AAW60037;

XX 11-SEP-1998 (first entry)

XX Antigenic peptide hFas (nd29) containing FC region.

XX Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
 KW heart failure; kidney failure; graft-versus-host disease; antibody;
 KW myocardial infarction; ischemic restenosis; endotoxic shock.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 1..16
 FT /note= "hFas antigen signal peptide"
 FT Protein 30..376

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:13:16 ; Search time 54.85 Seconds
(without alignments)
824.032 Million cell updates/sec

Title: US-09-742-454A-7
Perfect score: 1678
Sequence: 1 MARGSLRLRLRLVLGLWLA.....MHEALHNHYTKQSLSPCK 309

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	843	50.2	463	11 Q99LC4	Q99LC4 mus musculus
2	839	50.0	437	11 Q9RLA4	Q9RLA4 mus musculus
3	804.5	47.9	473	11 Q9D8L4	Q9D8L4 mus musculus
4	799	47.6	473	11 Q99L25	Q99L25 mus musculus
5	795.5	47.4	468	11 Q99L31	Q99L31 mus musculus
6	431.5	25.7	129	4 Q9NP84	Q9NP84 homo sapien
7	367	21.9	375	4 Q9BS21	Q9BS21 homo sapien
8	367	21.9	597	4 Q9BU10	Q9BU10 homo sapien
9	367	21.9	597	4 Q9BOB8	Q9BOB8 homo sapien
10	335.5	20.0	129	11 Q9QZK3	Q9QZK3 mus musculus
11	330.5	19.7	129	11 Q9CR75	Q9CR75 mus musculus
12	284.5	17.0	384	4 Q9UP60	Q9UP60 homo sapien
13	279.5	16.7	500	4 Q9BRV0	Q9BRV0 homo sapien
14	274.5	16.4	416	4 Q9NP66	Q9NP66 homo sapien
15	270	16.1	487	11 Q99KA4	Q99KA4 mus musculus
16	269.5	16.1	684	13 Q90544	Q90544 ginglymosto
17	269	16.0	426	11 Q9DCD9	Q9DCD9 mus musculus
18	255	15.2	479	11 Q99M22	Q99M22 mus musculus
19	255	15.2	484	11 Q99LA6	Q99LA6 mus musculus

20	197	11.7	94	4 Q9HCS0	Q9HCS0 homo sapien
21	185	11.0	268	13 Q90524	Q90524 ginglymosto
22	184.5	11.0	130	11 Q9D8W4	Q9D8W4 mus musculus
23	184	11.0	509	11 Q9QX57	Q9QX57 mus musculus
24	184	11.0	513	11 P97797	P97797 mus musculus
25	182.5	10.9	509	11 Q9WTN4	Q9WTN4 mus musculus
26	181	10.8	235	11 Q99M11	Q99M11 mus musculus
27	176.5	10.5	259	13 Q90530	Q90530 ginglymosto
28	176	10.5	506	6 Q46632	Q46632 bos taurus
29	173	10.3	509	11 Q08907	Q08907 mus musculus
30	172	10.3	506	6 Q46631	Q46631 bos taurus
31	171.5	10.2	257	13 Q90536	Q90536 ginglymosto
32	170	10.1	237	13 Q90545	Q90545 ginglymosto
33	168	10.0	105	11 Q99JC1	Q99JC1 mus musculus
34	167.5	10.0	252	13 Q90568	Q90568 ginglymosto
35	166.5	9.9	267	13 Q90529	Q90529 ginglymosto
36	165	9.8	238	7 Q9MXA2	Q9MXA2 aulonocara
37	162	9.7	261	7 Q19363	Q19363 sus scrofa
38	160	9.5	261	7 Q62868	Q62868 sus scrofa
39	157.5	9.4	237	7 Q9MX99	Q9MX99 aulonocara
40	155	9.2	261	7 Q98263	Q98263 sus scrofa
41	152.5	9.1	354	4 Q9NQK8	Q9NQK8 homo sapien
42	151	9.0	238	11 Q99M37	Q99M37 mus musculus
43	150	8.9	208	7 Q9MXA0	Q9MXA0 aulonocara
44	150	8.9	260	7 P79551	P79551 homo sapien
45	149.5	8.9	354	4 Q9P1W8	Q9P1W8 homo sapien

ALIGNMENTS

RESULT 1

Q99LC4 ID AC Q99LC4; PRELIMINARY; PRT; 463 AA.
AD Q99LC4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AAH03435.1; -
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match					50.2%; Score 843; DB 11; Length 463;
Best Local Similarity					46.8%; Pred. No. 2.3e-69;
Matches 161; Conservative 45; Mismatches 62; Indels 76; Gaps 6;					
Qy	24	SVAGEQAPGTAPCSRGSWSADLDKCMDCASCACRAPHSDCLGCAAP--PAPFRLWRS	81		
Db	138	SAAKTTPPSVYPLAPGSA-----AOTNSWTLGCLVKGYFPEPVVTWNS	182		
Qy	82	-----CDKTHT-----	87		
Db	183	GSLSGVHTFPAVLQSLDLYTLSSVTPSPSTWPTVTCNVAHPASSTKVKDKIIPRDCG	242		
Qy	88	CPP--CPAPEARGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE	145		
Db	243	CKPKICTVPEVS---SVFIPEPKPKDVLITLTTPKVTCTVVVDISKDDPEVQSFVDDVE	299		
Qy	146	VHNAKTPREEQYNSYRVSVTLVTHQDLNGKEVKCKVSNKALPAPTEKTSRAKGOP	205		
Db	300	VHTAQTPREEQFNSTFRSVSELPIMHQDLNGKEFKCKVNSAAPAPTEKTSIKTKGRP	359		
Qy	206	REPOVYTLPPSPREEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLDSDGS	265		

```
Db 360 KAPQVYTIPTPPKQMAKDKVSLTCMTIDFFPEDITVWQNGQPAENYKNTQPIMDTDGS 419
QY 266 FFYLSKLTVDKSWQGNVFCSCVMHEALHNHYTKSLSPCK 309
Db 420 YFYISKLNVQKSNWEAGNTFTCSVLHGLHNHHTKSLSPCK 463

RESULT 2
Q9R1A4
ID Q9R1A4 PRELIMINARY; PRT; 437 AA.
AC Q9R1A4;
DT 01-MAY-2000 (TREMELREL. 13, Created)
DT 01-MAY-2000 (TREMELREL. 13, Last sequence update)
DE 01-JUN-2001 (TREMELREL. 17, Last annotation update)
DE CAMVAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO IMMNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3E7D697C CRC64;

Query Match 50.0%; Score 839; DB 11; Length 437;
Best Local Similarity 46.5%; Pred. No. 5e-69;
Matches 160; Conservative 46; Mismatches 62; Indels 76; Gaps 6;

QY 24 SVAGEAQPAGTAPCSRGSWSADLDKCMDCASCRRPHSFCLGCAAP---PAPFRLLRWS 81
Db 112 SAAKTTPSVYPLAGSA-----AQTNSMVTGLGKVGYPPEPVTWNS 156
QY 82 -----CDKTH----- 87
Db 157 GLSGSGVHTFPVQLSDLYTLSSSVTPSPSEVTCTNVAHPASSTKVYDKKIVPRDCG 216
QY 88 CPD--CPAPEAGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGE 145
Db 217 CKPCICTVPEVS---SVFIFPPKPKDVLITLTPKTCVVVDISKDDPEVQFSWFVDVDE 273
QY 146 VHNAKTPREQNSTYRVVSVLTVLHQDLNGLNGKEYCKVSKNKAALPAPIEKTISKAGQP 205
Db 274 VHTAQTPREQNSTFRSVSELPIMHQDLNGLNGKEYCKVSKNKAALPAPIEKTISKGRP 333
QY 206 REPOVYTLPSREEMTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYKTPPVLDSDGS 265
Db 334 KAPQVYTIPTPPKQMAKDKVSLTCMTIDFFPEDITVWQNGQPAENYKNTQPIMDTDGS 393
QY 266 FFYLSKLTVDKSWQGNVFCSCVMHEALHNHYTKSLSPCK 309
Db 394 YFYISKLNVQKSNWEAGNTFTCSVLHGLHNHHTKSLSPCK 437

RESULT 3
Q9D8L4
```

```
ID Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4;
DT 01-JUN-2001 (TREMELREL. 17, Created)
DT 01-JUN-2001 (TREMELREL. 17, Last sequence update)
DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)
DE 1810060009RIK PROTEIN.
GN 1810060009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bojelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: TO IMMNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AK007918; BAB25349.1; -.
DR MGD; MGI:1924014; 1810060009RIK.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGV; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 47.9%; Score 804.5; DB 11; Length 473;
Best Local Similarity 53.1%; Pred. No. 8.1e-66;
Matches 154; Conservative 44; Mismatches 83; Indels 9; Gaps 3;

QY 20 ALLRSVAGEAQPAGTAPCSRGSWSADLDKCMDCASCRRPHSFCLGCAAPAPFRLRW 79
Db 193 ALLQS--GLYTLSSSVTVTNTWPSQTITCN-----VAHPASSTKVYDKKIEPRVP--ITQ 243
QY 80 RSDCKTHTCPCPAPEAGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
Db 244 NCPPLKECPCAAPDLLGGPSVFIFPKIKDVLMSLSPWTCVVVDVSEDDPVDVSW 303
QY 140 YVDGVEVHNKTPREQNSTYRVVSVLTVLHQDLNGLNGKEYCKVSKNKAALPAPIEKTIS 199
Db 304 FVNVVEVHTAQTOHREDYNTLRVVSALPIQHDNMSGKEFKCKVNNRALPSPIEKTIS 363
QY 200 KAGQPREPOVYTLPPSREEMTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYKTPPV 259
Db 364 KPRGPVRAPOVYVLPPLPPPAEEMTKKEFSLTCMTITGFLPAEIAVDWTSNGRTEQNTKNTATV 423
```


OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```
QY 268 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 309
DB 538 AHSILTVSEEWNTGETTYTCVVAHEALPNRVTERTVDKSTGK 579

RESULT 10
Q902W3 PRELIMINARY; PRT; 129 AA.
AC Q902W3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TYPE I TRANSMEMBRANE PROTEIN FN14.
GN FGFRP2 OR FN14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=20020297; PubMed=10551889;
RA Meighan-Wantha R.L., Hsu D.K.W., Guo Y., Brown S.A.N., Feng S.-L.Y.,
RA Peifley K.A., Alberts G.F., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Richards C.M., Winkles J.A.;
RT "The mitogen-inducible Fn14 gene encodes a type I transmembrane
RT protein that modulates fibroblast adhesion and migration.";
RL J. Biol. Chem. 274:33166-33176(1999).
DR EMBL; AF156164; AAF07882.1; -.
DR MGD; MGI:1351484; Fgfrp2.
KW Transmembrane.
SQ SEQUENCE 129 AA; 13637 MW; 14B5C68EEF493385 CRC64;

Query Match 20.0%; Score 335.5; DB 11; Length 129;
Best Local Similarity 55.1%; Pred. No. 2.4e-23;
Matches 70; Conservative 6; Mismatches 28; Indels 23; Gaps 2;

QY 1 MARGSLRLLRLVGLWIALRLSVAGEQAPGTPCSCGSSWSADLDKCMDCASCARPH 60
DB 1 MAPGPRSLPQLTVLGLVLMRAAAGEQAPGTPCSCGSSWSADLDKCMDCASCARPH 60
QY 61 SDFCLGCAAAPAPFRLI-----WRSCDKTHT-CPPCPAPEAE 97
DB 61 SDFCLGCAAAPAPFRLIWPILGGALSILVLLVSSFLVWRCRRRREKFTPIETGGE 120
QY 98 GAPSVFL 104
DB 121 GCPGVALL 127

RESULT 11
Q9CR75 PRELIMINARY; PRT; 129 AA.
AC Q9CR75;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FIBROBLAST GROWTH FACTOR REGULATED PROTEIN 2.
GN FGFRP2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA, AND CEREBELLUM;
RX MEDLINE=21083660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

QY 1 MARGSLRLLRLVGLWIALRLSVAGEQAPGTPCSCGSSWSADLDKCMDCASCARPH 60
DB 1 MAPGPRSLPQLTVLGLVLMRAAAGEQAPGTPCSCGSSWSADLDKCMDCASCARPH 60
QY 61 SDFCLGCAAAPAPFRLI-----WRSCDKTHT-CPPCPAPEAE 97
DB 61 SDFCLGCAAAPAPFRLIWPILGGALSILVLLVSSFLVWRCRRRREKFTPIETGGE 120
QY 98 GAPSVFL 104
DB 121 GCPGVALL 127

RESULT 12
Q9UP60 PRELIMINARY; PRT; 384 AA.
AC Q9UP60;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SNC73 PROTEIN.
GN SNC73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;
RT "Identification and characterization of SNC73, a gene which is down-
RT regulated in colorectal cancer.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF067420; AAC19365.1; -.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN.1.
SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 17.0%; Score 284.5; DB 4; Length 384;
Best Local Similarity 26.0%; Pred. No. 4.2e-18;
Matches 89; Conservative 38; Mismatches 138; Indels 77; Gaps 12;
```

```
QY 40 SWSADLDKCMDCASCRRPHSDFCLGCAAA---PPAPFRLWRS----- 81
Db 30 SSASPTSPKVFPLSLCSTQDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARFPSPQD 89
QY 82 -----CDKTHTC-----PPCPAPEAGAPSVFLFPPK 108
Db 90 ASGDLTYTSSQLTPATQCLAGKSVTKVHKYTNPSQDVTVPSPSTPPTSPST-PPT 148
QY 109 PK-----DTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKATKP 153
Db 149 PSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGL-RDASGVFTFTWTPSSGK--SAVQGP 205
QY 154 REEQYNSTYRVVSVLVTHQDNLNGKEYCKVSNKALPAPIETISKAKQPREPOVYTL 213
Db 206 PERDLGCGYSVSVLPCCAEFPWNGKFTCTAAAYPSKTPLTATLSKS-GNTRFPEVHLL 264
QY 214 PPSREEMTKNQ-VSLTCLVKGYFSPDSIAVEWESNGQ--PENNYKTTTPPVLD-SDG--SFF 267
Db 265 PPPSEELALNELVTLCLARGFSFKDVLVRLQGSQELPREKYLTVASRQEPGSGTTF 324
QY 268 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTKSLSPGK 309
Db 325 VTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGK 366

RESULT 13
Q9BRV0 PRELIMINARY; PRT; 500 AA.
AC Q9BRV0;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:14588).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005951; AAH05951.1; -.
SQ SEQUENCE 500 AA; 54154 MW; 0A9BF43F2A3CC6D9 CRC64;

Query Match 16.7%; Score 279.5; DB 4; Length 500;
Best Local Similarity 25.4%; Pred. No. 1.6e-17;
Matches 87; Conservative 41; Mismatches 137; Indels 77; Gaps 12;

QY 40 SWSADLDKCMDCASCRRPHSDFCLGCAAA---PPAPFRLWRS----- 81
Db 146 SSASPTSPKVFPLSLCSTQDGNVVIACLVQGFPOEPLSVTWSESGQGVTAARFPSPQD 205
QY 82 -----CDKTHTC-----PPCPAPEAGAPSVFLFPPK 108
Db 206 ASGDLTYTSSQLTPATQCLAGKSVTKVHKYTNPSQDVTVPSPSTPPTSPST-PPT 264
QY 109 PK-----DTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKATKP 153
Db 265 PSPSCCHPRLSLHRPALEDLLGSEANLTCTLTGL-RDASGVFTFTWTPSSGK--SAVQGP 321
QY 154 REEQYNSTYRVVSVLVTHQDNLNGKEYCKVSNKALPAPIETISKAKQPREPOVYTL 213
Db 322 PDRDLGCGYSVSVLPCCAEFPWNGKFTCTAAAYPSKTPLTATLSKS-GNTRFPEVHLL 380
QY 214 PPSREEMTKNQ-VSLTCLVKGYFSPDSIAVEWESNGQ--PENNYKTTTPPVLD-SDG--SFF 267
Db 381 PPPSEELALNELVTLCLARGFSFKDVLVRLQGSQELPREKYLTVASRQEPGSGTTF 440
QY 268 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTKSLSPGK 309
```

```
Db 441 VTSILRVAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGK 482

RESULT 14
Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RL "The European IMAGE consortium for integrated Molecular analysis of
RT human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AL389978; CAB97534.1; -.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 16.4%; Score 274.5; DB 4; Length 416;
Best Local Similarity 29.6%; Pred. No. 3.8e-17;
Matches 81; Conservative 38; Mismatches 116; Indels 39; Gaps 12;

QY 48 KCMDC--CASCRRAR---PHSDFCLGCAAAAPPAPERLLWRSCDKTHTCPCCPAPEAGCAPS 101
Db 152 QCPDGKSVTCHVKHYTNPSQDVTVPSPSTP-----PPCHPR----- 189
QY 102 VFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKATKPREEQYNST 161
Db 190 LSLHRPALED-LLLGSEANLTCTLTGL-RDASGATFTWTPSSGK--SAVQGPPELDLGGC 245
QY 162 YRVSVLTUHQDNLNGKEYCKVSNKALPAPIETISKAKQPREPOVYTLPPSREMT 221
Db 246 YSVSVLPVCAQPNWNGHGETFTCTAAHPKLTPLTANITKS-GNTRFPEVHLLPPPSSELA 304
QY 222 KNO-VSLTCLVKGYFSPDSIAVEWESNGQ--PENNYKTTTPPVLD-SDG--SFELYSKLTVD 275
Db 305 LNELVTLCLARGFSFKDVLVRLQGSQELPREKYLTVASRQEPGSGTTFVAVTSILRVA 364
QY 276 KSRWQGNVFCFSVMHEALHNHYTKSLSPGK 309
Db 365 AEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGK 398

RESULT 15
Q99KA4 PRELIMINARY; PRT; 487 AA.
AC Q99KA4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE UNKNOWN (PROTEIN FOR MGC:6727).
```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004786; AAH04786.1; -
SQ SEQUENCE 487 AA; 52554 MW; 7DC8E96DB333077B CRC64;

Query Match      16.1%; Score 270; DB 11; Length 487;
Best Local Similarity 30.6%; Pred. No. 1.2e-16;
Matches 77; Conservative 41; Mismatches 100; Indels 34; Gaps 11;

Qy 67 CAAPAPDFLLWRSCDKTHTCPCPAPEAGAPSVFLFPKPKDITLMISRTPEVTCVV 126
Db 243 CSGPPP-----PCPPCP-PSCH--PSLSLQRPALD-LLLGSDASLTCTLN 284
Qy 127 DVSHEDPEVKFNWYVDGVEVINAKTKPREQYNST---YRVVSULTVLHQDLNGKEYKC 183
Db 285 GLRNPEGAV-FTW-----EPSTGKDAVQKKAQVNSCGCYSSVSLPGCAERNWNSGAFKC 338
Qy 184 KVSNKALPAPTEKITSKAGQPREPOVYTLPPSREEMTKNO-VSLTCLVKGFPSPDIAVE 242
Db 339 TVTHPESDT-LTGTLAKITVNTFPQVHLPPPEELALNELVSLTCLVRAFNKEVLVR 397
Qy 243 WESNGQ---PENNYKTTTPPVLDSDG--SFFLYSKLTVDKSRWQGNVFCSCVMHEALHNH 297
Db 398 WLRGNEELSPESYLVFEPKPEGEGATTYLVTSVLRVSAETWKQGDQYSCMVGHEALPMN 457
Qy 298 YTKSLSLSPCK 309
Db 458 FTQKTIDRLSGK 469

```

Search completed: March 11, 2002, 15:23:24
Job time: 608 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:16 ; Search time 19.78 Seconds
(without alignments)
572.772 Million cell updates/sec

Title: US-09-742-454A-7

Perfect score: 1678

Sequence: 1 MARGSLRLRLVLGLWLA.....MHEALHNHYTKSLSLSPGK 309

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1224	72.9	330	1 GC1_HUMAN	P01857 homo sapien
2	1159.5	69.1	326	1 GC2_HUMAN	P01859 homo sapien
3	1151.5	68.6	290	1 GC3_HUMAN	P01860 homo sapien
4	1139	67.9	327	1 GC4_HUMAN	P01861 homo sapien
5	910.5	54.3	323	1 GC_RABIT	P01870 oryctolagus
6	876	52.2	329	1 GC2_CAVPO	P01862 cavia porce
7	838.5	50.0	324	1 GC1_MOUSE	P01868 mus musculus
8	838	49.9	326	1 GC1_RAT	P020759 rattus norv
9	836	49.8	329	1 GC3_MOUSE	P22436 mus musculus
10	833.5	49.7	393	1 GC1_MOUSE	P01869 mus musculus
11	825	49.2	333	1 GCB_RAT	P20761 rattus norv
12	825	49.2	398	1 GC3M_MOUSE	P03987 mus musculus
13	817	48.7	329	1 GCC_RAT	P20762 rattus norv
14	804.5	47.9	322	1 GCA_RAT	P20760 rattus norv
15	799.5	47.6	330	1 GCA_MOUSE	P01863 mus musculus
16	794.5	47.3	335	1 GCAB_MOUSE	P01864 mus musculus
17	794.5	47.3	399	1 GCAM_MOUSE	P01865 mus musculus
18	787	46.9	336	1 GCB_MOUSE	P01866 mus musculus
19	782	46.6	405	1 GCB_MOUSE	P01867 mus musculus
20	367	21.9	454	1 MUC_HUMAN	P01871 homo sapien
21	366.5	21.8	421	1 EPC_MOUSE	P06336 mus musculus
22	365	21.8	391	1 MUCB_HUMAN	P04220 homo sapien
23	360.5	21.5	429	1 EPC_RAT	P01855 rattus norv
24	353.5	21.1	428	1 EPC_HUMAN	P01854 homo sapien
25	349	20.8	455	1 MUC_MOUSE	P01872 mus musculus
26	344	20.5	458	1 MUC_RABIT	P03988 oryctolagus
27	339	20.2	476	1 MUC_MOUSE	P01873 mus musculus
28	336	20.0	454	1 MUC_MESAU	P06337 mesocricetu
29	334	19.9	479	1 MUCM_RABIT	P04221 oryctolagus
30	333	19.8	457	1 MUC_SUNMU	P20768 suncus muri
31	331.5	19.8	450	1 MUC_CANFA	P01874 canis fami
32	302.5	18.0	299	1 ALC_RABIT	P01879 oryctolagus
33	290	17.3	438	1 HVC2_HETER	P23085 heterodontu

34	288.5	17.2	446	1 MUC_CHICK	P01875 gallus gall
35	284.5	17.0	438	1 HVC5_HETER	P23087 heterodontu
36	282	16.8	353	1 ALC1_HUMAN	P01876 homo sapien
37	279	16.6	353	1 ALC1_GORGO	P20758 gorilla gor
38	275.5	16.4	461	1 HVCM_HETER	P23088 heterodontu
39	274.5	16.4	340	1 ALC2_HUMAN	P01877 homo sapien
40	274	16.3	370	1 HVC1_HETER	P23084 heterodontu
41	273.5	16.3	393	1 HVC3_HETER	P23086 heterodontu
42	255	15.2	344	1 ALC_MOUSE	P01878 mus musculus
43	252.5	15.0	481	1 MUCM_ICTPU	P23735 ictalurus p
44	178.5	10.6	105	1 LAC1_MOUSE	P01843 mus musculus
45	173	10.3	104	1 LAC2_RAT	P20767 rattus norv

ALIGNMENTS

RESULT 1					
ID	GC1_HUMAN	STANDARD;	PRT;	330 AA.	
AC	P01857;	1986 (Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	15-JUL-1999	(Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.				
GN	IGHG1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=82274238; PubMed=6287432;				
RA	Ellison J.W., Berson B.J., Hood L.E.;				
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."				
RL	Nucleic Acids Res. 10:4071-4079(1982).				
RP	[2]				
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).				
RX	MEDLINE=71064024; PubMed=5489771;				
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,				
RT	Waxdal M.J., Edelman G.M.;				
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino				
RL	acid sequence of heavy-chain cyanogen bromide fragments H1-H4."				
RN	Biochemistry 9:3161-3170(1970).				
RP	[3]				
RP	SEQUENCE OF 136-329 (EU).				
RX	MEDLINE=71064025; PubMed=5530842;				
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,				
RT	Edelman G.M.;				
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino				
RL	acid sequence of heavy-chain cyanogen bromide fragments H5-H7."				
RN	Biochemistry 9:3171-3181(1970).				
RP	[4]				
RP	SEQUENCE (MYELOMA PROTEIN NIE).				
RX	MEDLINE=77070269; PubMed=826475;				
RA	Ponstingl H., Hilschmann N.;				
RT	"The rule of antibody structure. The primary structure of a				
RT	monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The				
RT	chymotryptic peptides of the H-chain, alignment of the tryptic				
RT	peptides and discussion of the complete structure."				
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).				
RN	[5]				
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.				
RX	MEDLINE=83289131; PubMed=6884994;				
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;				
RT	"Three-dimensional structure determination of antibodies. Primary				
RT	structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."				
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).				
RN	[6]				
RP	DISULFIDE BONDS.				
RX	MEDLINE=71064027; PubMed=4923144;				
RA	Gall W.E., Edelman G.M.;				
RT	"The covalent structure of a human gamma G-immunoglobulin. X.				

[2]
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
 RX MEDLINE=81007873; PubMed=6774012;
 RA Wang A.-C., Tung E., Fudenberg H.H.;
 RT "The primary structure of a human IgG2 heavy chain: genetic,
 RL evolutionary, and functional implications.";
 RN J. Immunol. 125:1048-1054(1980).
 [3]
 RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
 RX MEDLINE=80001357; PubMed=113060;
 RA Connell G.E., Parr D.M., Hofmann T.;
 RT "The amino acid sequences of the three heavy chain constant region
 RL domains of a human IgG2 myeloma protein.";
 RN Can. J. Biochem. 57:758-767(1979).
 [4]
 RP SEQUENCE OF 238-275 (ZIE).
 RX MEDLINE=80114419; PubMed=118920;
 RA Hofmann T., Parr D.M.;
 RT "A note of the amino acid sequence of residues 381-391 of human
 RL immunoglobulins gamma chains";
 RN Mol. Immunol. 16:923-925(1979).
 [5]
 RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
 RX Hofmann T., Parr D.M.;
 RL Submitted (MAR-1980) to the PIR data bank.
 [6]
 RP SEQUENCE OF 1-121 (DOT).
 RX MEDLINE=9525298; PubMed=7737190;
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
 RT "Characterization of the two unique human anti-flavin monoclonal
 RL immunoglobulins.";
 RN Eur. J. Biochem. 228:886-893(1995).
 [7]
 RP DISULFIDE BONDS.
 RX MEDLINE=72033500; PubMed=4940472;
 RA Milstein C., Frangione B.;
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
 RN Biochem. J. 121:217-225(1971).
 [8]
 RP DISULFIDE BONDS.
 RX MEDLINE=69064124; PubMed=5782707;
 RA Frangione B., Milstein C., Pink J.R.L.;
 RT "Structural studies of immunoglobulin G.";
 RN Nature 221:145-148(1969).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; V00554; CAB58438.1; -;
 DR PIR; A02148; G2HU.
 DR MIM; 147110; -;
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig-cl.
 DR InterPro; IPR003600; Ig-like.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00407; Igcl; 2.
 DR SMART; SM00410; Ig-like; 1.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 140 200
 FT DISULFID 246 304
 FT SITE 156 156
 FT MOD_RES 326 326
 FT VARIANT 60 60
 FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;
 SQ
 Query Match 69.1%; Score 1159.5; DB 1; Length 326;
 Best Local Similarity 66.6%; Pred. No. 6.6e-82;
 Matches 229; Conservative 8; Mismatches 30; Indels 77; Gaps 5;
 Oy 24 SVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRAPHSDFCLGCAAP--PAPFRLWRS 81
 Db 2 STKGPSVFPPLAPCSRSTS-----ESTAAIGCLVKDYFFPEPVTWSNS 43
 Oy 82 -----CDKTH----- 86
 Db 44 GALTSGVHTFFAVLQSSGLYSLSSVTVPSNFGTQTYTCNVDDHKPSNTKVDKTVVERKCC 103
 Oy 87 -TCPPCPAPEAGAPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWVDGVE 145
 Db 104 VECPCPAPPVAG-PSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVQFNWVDGVE 162
 Oy 146 VHNATKPREQYNSTYRVSVLTVLDHQLNGKYEKCKVSKNKPAPLPIETISAKGQP 205
 Db 163 VHNATKPREQFNSTFRVSVLTVLDHQLNGKYEKCKVSKNKPAPLPIETISAKGQP 222
 Oy 206 REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 265
 Db 223 REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 282
 Oy 266 FFLYSKLVDSRWQGNVFCSCVMHEALHNHYTKSLSLSPGK 309
 Db 283 FFLYSKLVDSRWQGNVFCSCVMHEALHNHYTKSLSLSPGK 326
 RESULT 3
 GC3_HUMAN
 ID GC3_HUMAN STANDARD; PRT; 290 AA.
 AC P01860;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).
 GN IGHG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE (DISEASE PROTEIN WIS).
 RX MEDLINE=81021548; PubMed=6774747;
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
 RL gamma 3 heavy-chain disease protein Wis.";
 RN Biochemistry 19:4304-4308(1980).
 [2]
 RP REVISIONS TO 12-97 OF PROTEIN WIS.
 RX MEDLINE=77118561; PubMed=402363;
 RA Michaelson T.E., Frangione B., Franklin E.C.;
 RT "Primary structure of the 'hinge' region of human IgG3. Probable
 RL quadruplication of a 15-amino acid residue basic unit.";
 RN J. Biol. Chem. 252:883-889(1977).
 [3]
 RP REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).
 RX MEDLINE=77021516; PubMed=823945;
 RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
 RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
 RL Structure of the Fc fragment of immunoglobulin G3.";


```

SEQUENCE OF I-128.
MEDLINE=76135469; PubMed=1243651;
Pratt D.M., Mole L.E.;
"Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin G of different allotype."; Biochem. J. 151:337-349(1975). [3]
SEQUENCE OF 88-266 FROM N.A. MEDLINE=83299917; PubMed=6193512; Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.; "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy chain and identification of two genomic C gamma genes."; Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982). [4]
SEQUENCE OF 132-161. MEDLINE=70110015; PubMed=5461106; Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.; "Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin G"; Biochem. J. 116:249-259(1970). [5]
SEQUENCE OF 129-131 AND 155-322. Hill R.B., Lebovitz H.E., Fellows R.E. Jr., Delaney R.; (in) Killander J. (eds.); Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell, Stockholm (1967). CC
-!- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER, 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15 MARKERS AND REF.5 THE E15 MARKER. CC
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch). CC
EMBL: M16426; AAA31289.1; -. DR
PIR: A02161; GHBB DR
InterPro: IPRO03006; IG_MHC. DR
InterPro: IPRO03597; IG-cl. DR
Pfam: PF00047; ig; 3. DR
SMART: SM00407; Igc1; 2. DR
PROSITE: PS00290; IG_MHC; 1. DR
Immunoglobulin domain; Immunoglobulin C region. KW
NON_TER 1 FT
VARIANT 104 104 T -> M (IN D11 MARKER). FT
VARIANT 185 185 T -> A (IN E15 MARKER). FT
CONFLICT 48 48 N -> E (IN REF. 2). FT
CONFLICT 71 71 V -> VPV (IN REF. 2). FT
CONFLICT 144 144 Q -> E (IN REF. 3 AND 4). FT
CONFLICT 173 173 N -> D (IN REF. 5). FT
CONFLICT 187 187 Q -> E (IN REF. 3 AND 5). FT
CONFLICT 201 201 N -> D (IN REF. 5). FT
CONFLICT 218 218 Q -> E (IN REF. 5). FT
CONFLICT 233 233 E -> Q (IN REF. 5). FT
CONFLICT 246 246 N -> D (IN REF. 5). FT
CONFLICT 256 256 E -> G (IN REF. 5). FT
CONFLICT 260 260 N -> D (IN REF. 5). FT
CONFLICT 266 266 N -> D (IN REF. 5). FT
CONFLICT 280 280 Y -> W (IN REF. 5). FT
CONFLICT 284 284 N -> S (IN REF. 5). FT
SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64; SQ
Query Match 54.3%; Score 910.5; DB 1; Length 323;
Best Local Similarity 52.7%; Pred. No. 7.6e-63;
Matches 178; Conservative 31; Mismatches 54; Indels 75; Gaps 6;
QY 29 QAPGT---APCSRGSSWSADLDKCMDCASCRAPHSDFCIGCAAAP--PAFPFLLRMS-- 81
::| | | | | | | | | | | | | | | | :
Db 4 KAPSVEPLACP-----CGDPSPSTVTLCGLVKGYLPPEPVTTWNSGT 45

```

QY 82 -----CDKTH-----TC--PPC 91
 Db 46 LTNGVTRFPVSQSGLYSLSSVSVTSQSPVTCNVAHPATNTKDKTVAPSTCSKPTC 105
 QY 92 PAPAEGAPSVFLPPPKKDTLMTSRPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKT 151
 Db 106 PPELLGGPSVFIIPPKKDTLMTSRPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKT 165
 QY 152 KPREQVNSYRVVSVLTVLHDLNKGKEYCKVSNKALPAPIEKTISKAKGQPREPQVY 211
 Db 166 PLRQEQNSTIRVVSTLPTHQDWLGRGKFKCKVHNKALPAPIEKTISKAKGQPLEPKVY 225
 QY 212 TLPSREEMTKNOVSLCLVKGFPYSDIAVWESNGOPENNYKTPPPVLDSDGSGFFLYSK 271
 Db 226 TMGPPRELSSRSVSLTCMNGFPPSDISVEWENKRAEDNYKTPPAVLSDGSGFFLYSK 285
 QY 272 LTVDKSRWQGNVSCSVMMHEALHNHYTQKSLSLSPGK 309
 Db 286 LSVPTSEWQGDVFTCSVMHEALHNHYTQKSLSRSPGK 323

RESULT 6

GC2_CAVPO ID GC2_CAVPO STANDARD; PRT; 329 AA.
 AC P01862;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-2 CHAIN C REGION.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE OF 1-3.
 RA Trischmann T.M.;
 RL Submitted (APR-1975) to the PIR data bank.
 RN [2]
 RP SEQUENCE OF 4-68.
 RX MEDLINE=71058471; PubMed=5538606;
 RA Birshlein B.K., Hussain Q.Z., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 immunoglobulin-G(2). 3. Amino acid sequence of the region around the
 half-cysteine joining heavy and light chains.";
 RL Biochemistry 10:18-25(1971).
 RN [3]
 RP SEQUENCE OF 69-133 AND 312-329.
 RX MEDLINE=71058486; PubMed=5538616;
 RA Turner K.J., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
 and hinge region cyanogen bromide fragments.";
 RL Biochemistry 10:9-17(1971).
 RN [4]
 RP SEQUENCE OF 134-226.
 RX MEDLINE=75036072; PubMed=4429665;
 RA Tracey D.E., Cebra J.J.;
 RT "Primary structure of the CH2 homology region from guinea pig IgG2
 antibodies.";
 RL Biochemistry 13:4796-4803(1974).
 RN [5]
 RP SEQUENCE OF 227-311.
 RX MEDLINE=75036073; PubMed=4609467;
 RA Trischmann T.M., Cebra J.J.;
 RT "Primary structure of the CH3 homology region from guinea pig IgG2
 antibodies.";
 RL Biochemistry 13:4804-4811(1974).
 RN [6]
 RP DISULFIDE BONDS.
 RX MEDLINE=71058474; PubMed=4922544;
 RA Oliveira B., Lamm M.E.;
 RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";

RL Biochemistry 10:26-31(1971).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
 CC 13 INBRED GUINEA PIGS.
 DR PIR: A02151; G2GP.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003597; Ig_GL.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 3.
 DR SMART: SM00407; IgCL; 2.
 DR SMART: SM00410; Ig_Like; 1.
 DR PROSITE: PS00290; Ig_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
 FT NON_TER 1
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 28 79
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 142 202
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
 FT DISULFID 248 308
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;
 Query Match 52.2%; Score 876; DB 1; Length 329;
 Best Local Similarity 54.1%; Pred. No. 3.4e-60;
 Matches 171; Conservative 28; Mismatches 57; Indels 60; Gaps 4;
 QY 53 ASCRARPHSDFCLGCAAP--PAPFRLWRS----- 81
 Db 14 ASCVDTSGMTLGLGVGFPEPTVKNWNSGALTSVHTFPVAVLQSGLYSLTSMVTVP 73
 QY 82 -----CDKTH-----TCPPCPAPEAGCAPSVFLPPKPKDTLM 114
 Db 74 SQKATCNVAHPASSTKVDKTVETPZPBCTCPKPPENLGGPSVFIIPPKKDTLM 133
 QY 115 ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKTPREQVNSYRVVSVLTVLHQD 174
 Db 134 ISLTPTVTCVVVDVSHEDPEVKFNWYVDGVVHNAKTPREQVNSYRVVSVLTVLHQD 193
 QY 175 WLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGF 234
 Db 194 WLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGF 253
 QY 235 YPSDIAVWESNGOP--ENNYKTPPPVLDSDGSGFFLYSKLTVDKSRWQQGNVSCSVH 292
 Db 254 FPADIHWASNRVPSSEKYNTPPIEDADGSGFFLYSKLTVDKSRWQQGNVSCSVH 313
 QY 293 ALNHYTQKSLSLSPG 308
 Db 314 ALNHYTQKSLSLSPG 329
 RESULT 7
 GC1_MOUSE ID GC1_MOUSE STANDARD; PRT; 324 AA.
 AC P01868;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE IG GAMMA-1 CHAIN C REGION.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 gamma 1 chain gene.";
 RL Cell 18:559-568(1979).

Best Local Similarity 50.0%; Pred. No. 2.5e-57;		Matches 163; Conservative 55; Mismatches 79; Indels 29; Gaps	
QY	1 MARGSLRLLRLVLGLWLALLRLRSVAGEQAPGTAPCSRGSWS-----ADLDKCM 50		
Db	11 LAPGSAATQNSWTLG---CLVKGYPF--PVTVTNWSGLSSGVHTFPVQLQSDLYTL 65		
QY	51 DCASCRARPHSDFCGLCAAAAPPAPF-----RLLRSCDKTHTCTPP--CPAPEAEGAPSVF 100		
Db	66 SSVTPSPSPRSETVTVCNAHPASSTVKDKKIVPRDCG----CKPCICTVPEVS---SVF 111		
QY	104 LFPKPKDTLMISPTPEVTCVVVDVSHEDPEVKFNWYGVGVEHNATKPREEOYNSTYR 160		
Db	119 IFPPKPKDVITITTPKVTCTVVDISKDDPEVQSFVDFVEVHTAQTQPREQFNSTER 177		
QY	164 VVSVLTVLHODWLNKKEYCKVSNKALPAPTEKTSKAGQPRPQVYVTLPPSPREEMTKN 220		
Db	179 SVSELPIMHODWLNKKEFKCRVNSAAPAPIEKTSIKGRKAPQVYVTLPPKQMAKD 236		
QY	224 QVSLTCLVKGYFSDIAVENESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGN 280		
Db	239 KVSLTCTMITDFEFDIEVQWQNPQAEYKNTQPIIMNTNGSYEVYSKLVQKNSWEAGN 298		
QY	284 VFCSVMHEALHNHYTOKSLSLSPGK 309		
Db	299 TFCSVLHEGLHNHTEKSLSHSPGK 324		
RESULT 8			
ICL_RATE	GC_L_RATE	STANDARD;	PRT; 326 AA.
AC	P20759;		
DT	01-FEB-1991 (Rel. 17, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DE	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	IG GAMMA-1 CHAIN C REGION.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
ON	NCBI_TaxID=10116;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=89232738; PubMed=3149946;		
RA	Brueggemann M.		
RA	"Evolution of the rat immunoglobulin gamma heavy-chain gene family."		
RL	Gene 74:473-482(1988).		
DR	PIR; PS0017; PS0017.		
DR	InterPro; IPR003006; Iq_MHC.		
DR	InterPro; IPR003597; Iq_cl.		
DR	Pfam; PF00047; Iq; 3.		
DR	SMART; SM00407; IGcl; 2.		
DR	PROSITE; PS00290; IG_MHC; 1.		
KW	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.		
FT	NON_TER 1 97		CHI.
FT	DOMAIN 1 97		HINGE.
FT	DOMAIN 98 112		CH2.
FT	DOMAIN 113 219		CH3.
FT	DOMAIN 220 326		
FT	DISULFID 27 82		
FT	DISULFID 102 102		INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 106 106		INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 109 109		INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 111 111		INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 140 200		
FT	DISULFID 246 304		
FT	CARBOHYD 176 176		
QY	SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;		
Query Match 49.9%; Score 838; DB 1; Length 326;			
Best Local Similarity 60.5%; Pred. No. 2.8e-57;			
Matches 150; Conservative 43; Mismatches 47; Indels 8; Gaps			

```
Qy 67 CAAAPPAPF-----RLLRSCDKTHTCPAPAEAGAPSVFLPPPKDKTLMISRTPEV 121
Db 82 CNVAHPASSTVADKKIYPRNGC--GDKPCICTGSE--VSSVFIFPPPKDVLTTLPKV 138
Qy 122 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVSLTVLHODWLNKGEY 181
Db 139 TCVVVDISQDDPEVHFVSWFVDNKEVHTAWTPQAEQYNSTFRVVSALPIQHDMMRGKE 200
Qy 182 CKKYSNKALPAPIETKTSKAGQPREQVYTLPPSRREMTKNQVSLCLVKGFPSPDIAY 241
Db 199 RCKYTSAAFPSPIEKTIKSPGRTQVPHVYTPMTSKTEMTQNEYSITCMWVGFPDPDIY 258
Qy 242 EWESNGQPNENKYPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQ 300
Db 259 EWQNGQPNENKYPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQ 320
Qy 302 SLSLSPGK 309
Db 319 SLSHSPGK 326
```

RESULT 9

```
GC3_MOUSE ID GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; -, NOT_ANNOTATED_CDS.
DR PIR; B02156; G3WSC.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00410; IG.Like; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;
```

```
Query Match 49.8%; Score 836; DB 1; Length 329;
Best Local Similarity 62.7%; Pred. No. 4e-57;
Matches 156; Conservative 33; Mismatches 54; Indels 6; Gaps 2;
```

```
Qy 67 CAAAPPAPFRLLRSCD-----KTHTCP--PCPAEAGAPSVFLPPPKDKTLMISRTPE 120
```

```
Db 81 CNVAHPASKTELKRIEPRIPKSPTPGSCPPGNGILGPSVFIPPKPKDMLSLTPK 140
Qy 121 VTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVSLTVLHODWLNKGE 180
Db 141 VTCVVVDVSEDDPDVHVSFWVDNKEVHTAWTPQAEQYNSTFRVVSALPIQHDMMRGKE 200
Qy 181 YKCVSNKALPAPIETKTSKAGQPREQVYTLPPSRREMTKNQVSLCLVKGFPSPDIAY 240
Db 201 FKCKVNNKALPAPIETLSKPKGQAQTPQVYTIPTPREQMSKKVSLTCLTVTFSEALS 260
Qy 241 VEWESNGQPNENKYPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQ 300
Db 261 VEWERNCELEQDYKNTPPILSDGTYFLYSLKLTVDTSMLQGEIFTCVSVHEALHNHYTQ 320
Qy 301 KSLSLSPGK 309
Db 321 KNLRSRSPGK 329
```

RESULT 10

```
GC1M_MOUSE ID GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
RT conserved transmembrane sequence and a 28-residue intracellular
RT domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl terminl of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; G1MSM.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_c1.
DR SMART; SM00407; Ig_c1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 INTERCHAIN (WITH A HEAVY CHAIN).
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
FT SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;
```

```
Query Match 49.7%; Score 833.5; DB 1; Length 393;
Best Local Similarity 49.8%; Pred. No. 7.6e-57;
Matches 162; Conservative 55; Mismatches 79; Indels 29; Gaps 7;

QY 1 MARGSLRLLRLVLLGLWALRLSVAGEAQTAPCSRGSWS-----ADLDKCM 50
DB 11 LAPGSAATNSMTLIG---CLVKGYPE--PVTVTWNSGLSSGVHTFFAVLQSDLYTIS 65

QY 51 DCASCRARPHSDFCLGCAAPPAPE-----RLLRSCDKTHRCPP--CPAPAEAGPSVF 103
DB 66 SSVTPSSRPSETVTCNVAPASSTKVKIYPRDCG----CKPICIVPEVS---SVF 118

QY 104 LFPKPKDLMISRTEVTCVVDVSHEDPEVKFNWYDGVVEYHNAKTKPREQYNSTYR 163
DB 119 IFPPKPKDVLITLTPKVTGVVDVSDISKDDPEVQFSWFVDVVEHTAQTQPREQFNSTFR 178

QY 164 VVSVLTVLHODWLNKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREMTKN 223
DB 179 SVSELPIMHQDLNGKEFKRCYNSAFAFPAPIEKTISKTKGRKAPQVYTIPTPKQMAKD 238

QY 224 QVSLCTLVKGFPSDIAVESWESNGPENNYKTPPVLDSGDFELYSKLTVDKSRWQGN 283
DB 239 KVSILCTWITDFPEDITVQWNGQPAENYKNTQPIIMNTNGSYFYVSKLVNOKRSNWEAGN 298

QY 284 VFSCSYMEALHNHYTKSLSPG 308
DB 299 TITCSVLHESLHNHHTKSLSHSPG 323

RESULT 11
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761.
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2B CHAIN C REGION.
OS Rattus norvegicus (Rat).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0018; PS0018.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Ig_c1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 49.2%; Score 825; DB 1; Length 333;
Best Local Similarity 64.0%; Pred. No. 2.8e-56;
Matches 146; Conservative 33; Mismatches 49; Indels 0; Gaps 0;

QY 82 CDKTHTCPPCPAPEAGAPSVFLFPKPKDTLMISRTEVTCVVDVSHEDPEVKFNWY 141
DB 106 CPTCPTCHKCPVPELLGGPSVFIFFPKPKDILLISQNAKVTCTVVDVSEEDPVQFSWF 165

QY 142 DGVVEHNAKTKPREQYNSTYRVVSVLTVLHODWLNKEKCKVSNKALPAPIEKTISKA 201
DB 166 NNVEHTAQTQPREQYNSTYRVVSALPIQHQDMSCGKFKVNNKALPSPIEKTISK 225

QY 202 KGPREPOVYTLPPSREMTKNQVSLTCLVGFYPSDIAVESWESNGPENNYKTPPVL 261
DB 226 KGLVRPQVYVMGPPTTEQLTEQVSLTCLSGFLPNDIGVETWSNGHIEKNYKNTPE 285

QY 262 SDGSFELYSLKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSPG 309
DB 286 SDGSFEMYSKLVNRSRWSRAPFVCSVHGLNHHVKSISRPPCK 333

RESULT 12
GCB_MOUSE
ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987.
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A.; Word C.J.; Rimm D.; Der-Balan G.P.; Martinez H.M.;
RA Tucker P.W.; Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
```

RA Komatomy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
RT segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00451; AAB59655.1; -;
DR EMBL: V01526; CAA24767.1; ALT_SEQ.
DR PIR: A02155; G3MSM.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igcl; 2.
DR SMART: SM00410; Ig_like; 1.
DR PROSITE: PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7E264B50A41B95 CRC64;

Query Match 49.2%; Score 825; DB 1; Length 398;
Best Local Similarity 62.3%; Pred. No. 3.4e-56;
Matches 154; Conservative 33; Mismatches 54; Indels 6; Gaps 2;

QY 67 CAAPAPPAPFRLMLRSCD-----KTHTCP--PCPAEAGAPSVFLPPKPKDTLMISRTPE 120
DB 81 CNAHPASKTELKRIEPRPKSTPPGSSCPPGNIILGSPVFIFPPKPKDALMISLTPK 140
QY 121 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKE 180
DB 141 VTCVVVDVSEDDPDVHVSWFVFNKVEVHTAWTPQREAQYNSTFVVVSAALPIQHQQDMMRGKE 200
QY 181 YKCVSNKALPAPIEKTISKAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIA 240
DB 201 FKCKVNNKALPAPIERTISKRAQTPQVYTIIPPREQMSKKKVSITCLVTFNFFSEAIS 260
QY 241 VEWESNGPENNYKTTTPVLDSGSEFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQ 300
DB 261 VEWERNELQDYKNTPIPLDSGTFFLYSKLTVDTSLWLGQEIFTCVSVHEALHNHHTQ 320
QY 301 KSLSLSP 307
DB 321 KNLRSRP 327

RESULT 13
GCC_RAT
ID GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2C CHAIN C REGION.
OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_RaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RT region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07189; CAA30169.1; -;
DR PIR: S00847; S00847.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igcl; 2.
DR SMART: SM00410; Ig_like; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 48.7%; Score 817; DB 1; Length 329;
Best Local Similarity 53.8%; Pred. No. 1.1e-55;
Matches 155; Conservative 46; Mismatches 49; Indels 38; Gaps 3;

QY 33 TAPCSRSSWS-----ADLDKMDCCASCARPHSDFCGLGCAAPAPFRLMLRS 81
DB 69 TVPSSTWSSQTVCVAHPATKSNLIKRIEPRPKPRPTDIC-----S 112
QY 82 CDKTHTCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 141
DB 113 CD-----DNLGRPSVFIFFPKPKDTLMTLPKVTVCVVVDVSEEDPVQFSWFV 161
QY 142 DGEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKA 201
DB 162 DNVKFTAQTPQHEEQUNGTFRVVYSTLHIHQDMMWSGKEFKCKVNNKDLPSPIEKTISK 221
QY 202 KGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGPENNYKTTTPPVLD 261
DB 222 RGKARTQVYTIIPPREQMSKKKVSITCLVTFNFFSEAISVEMERNGELEQDYKNTLPVLD 281
QY 262 SDGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKLSLSPGK 309
DB 282 SDESIFLYSKLSDTDSWMRGDIYTCVSVHEALHNHHTQKNLSRSPGK 329

RESULT 14
GCC_RAT
ID GCC_RAT STANDARD; PRT; 322 AA.
AC P20760;
DT 01-FEB-1991 (Rel. 17, Created)


```

DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; E8EAL36A9DE01EDB CRC64;

Query Match 47.9%; Score 804.5; DB 1; Length 322;
Best Local Similarity 47.0%; Pred. No. 1e-54;
Matches 157; Conservative 38; Mismatches 62; Indels 77; Gaps 5;

Qy 30 APGTAPCRGSSWADLCKMDCASCRRPHSDFCGLCAAP--PAPFRLWRS----- 81
||||| :| ||| ||| :| |
Dy 12 APGTA-----LKNSMVTGLCLVKGYFPEPVTVTVNSGALSSG 49
||||| :| ||| ||| :| |
Qy 82 -----CDKTH-----TCPPCPAPE 95
||||| :| ||| ||| :| |
Dy 50 VHTFPVQLQGLYLTSSVTPSSWSSQAVTCNVHPASSTKYVDKKIVPRECNPCGCTG 109
||||| :| ||| ||| :| |
Qy 96 AEGAPSVFLFPKPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNKTKPRE 155
:| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 110 SE-VSSVFIFPPKTKDVLITLTKPVTCVVVDISQNDPEVFESWFDIDVEVHTAQTSHAPE 168
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Qy 156 EQYNSTYRVSVLVFVLHODWLNKEIKYCKVSNKALPAPIEKTIISKAKQPREPVYITLPP 215
:| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 169 KOSNLTSLRSVSELPVHRLDNLNGTKFKCKVNSGAPPAPPIEKISKEPTGPRGPQYITMAP 228
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Qy 216 SREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPTPVLDSDGSFSLSKLTVD 275
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 229 PREMTQSQVSLTCLVKGFYPPDIYETWMKNGQPEQNTKPTPTMDTGSLYSLKLVNK 288
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Qy 276 KSRWQGNVFCVSVMEALHNHYTKQSLSPGK 309
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|
Dy 289 KETWQGNFTFCVSLHLEGLNHHTEKSLSPGK 322
||||| :| ||| ||| :| ||| ||| :| ||| ||| :| ||| ||| :|

RESULT 15
GCAA_MOUSE STANDARD; PRT; 330 AA.
ID GCAA_MOUSE

```

```

AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RL Nucleic Acids Res. 8:3143-3155(1980).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
RT and evolution of heavy chain genes: further evidence for intervening
RT sequence-mediated domain transfer.";
RL Nucleic Acids Res. 9:1365-1381(1981).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Olio R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN (4)
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin:amino-acid sequence of the Fc fragment. Implications
RT for the evolution of immunoglobulin structure and function.";
RL Eur. J. Biochem. 43:423-435(1974).
RN (5)
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
RL immunoglobulin. Identification of the disulfide bridges.";
RL Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1; -
DR PIR; A02152; G2MSA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igcl; 2.
DR SMART; SM00410; Ig_like; 1.
DR PROSITE; PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204

```

```
FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 47.6%; Score 799.5; DB 1; Length 330;
Best Local Similarity 49.7%; Pred. No. 2.5e-54;
Matches 157; Conservative 29; Mismatches 69; Indels 61; Gaps 4;

QY 55 CRARPHSDCLGCAAP--PAPERLLWRS----- 81
Db 15 CGDTTGSSTLGLVKGYPPEVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVTSST 74
QY 82 -----CDKTH-----TCPP--CPAPEAEGAPSVFLFPPKPKDYL 113
Db 75 WPSQSTCNVAHASPSTKVDKIEPRGPTIKPCPPCKPAPNLLGGPSVIFPPKIDVL 134
QY 114 MISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQ 173
Db 135 MISLPIVTCVVVDVSEDDPDVQISWFEVNNVEVHTAQTQTHREDYNSTLRVVSALPIQH 194
QY 174 DWLNGKEYCKVKNKALPADIETISKAKQCPREPQVYYTLPPSREEMTKNQVSLTCLVK 233
Db 195 DMSGKEFKCKVNNKDLPAIERTISKPGSVRAPQYIVLPPPEEEMTKKQVTLTCMVD 254
QY 234 FYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVNHEA 293
Db 255 FMPEDIVETWNTNGKTELNKTPEVLDSDGSYFMYSKLRVEKKNWVERNSYSCSVVHEG 314
QY 294 LHNHYTKSLSPGK 309
Db 315 LNHHTTKSFRTPGK 330
```

Search completed: March 11, 2002, 15:23:50
Job time: 574 sec

C;Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C;Accession: S31866

R;Filpula, D.
submitted to the EMBL Data Library, February 1993
A;Description: Screening method for protein-protein interactions of cloned gene products.
A;Reference number: S31866
A;Accession: S31866
A;Molecule type: mRNA
A;Residues: 1-255 <FIL>
A;Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069
C;Keywords: Immunoglobulin
F;1-22/Region: Escherichia coli outer membrane protein A precursor
F;23-255/Region: human Ig gamma-1 chain C region

Query Match 72.9%; Score 1224; DB 4; Length 255;
Best Local Similarity 97.4%; Pred. No. 4e-81;
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHTCPCPAEAGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 139
Db :|||||
26 KSCDKTHTCPCPAELGPGSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 85
QY 140 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
Db :|||||
86 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 145
QY 200 KAKGPQREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 259
Db :|||||
146 KAKGPQREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 205
QY 260 LDSGSEFLYSLKTVDSKRWQGVNFVSCVYHVALHNYTKQSLSLSPGK 309
Db :|||||
206 LDSGSEFLYSLKTVDSKRWQGVNFVSCVYHVALHNYTKQSLSLSPGK 255

RESULT 3

IGHU
Ig gamma-1 chain C region - human

C;Species: Homo sapiens (man)
C;Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.
Nucleic Acids Res. 10, 4071-4079, 1982
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.
A;Reference number: A93433; MUID:82274238
A;Accession: A93433

A;Molecule type: DNA
A;Residues: 1-330 <ELL>
A;Cross-references: EMBL:Z17370
A;Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) markers.
A;Note: Lys-330 is removed after translation
R;Harris, L.J.
submitted to the EMBL Data Library, October 1992
A;Reference number: S33904
A;Accession: S36861
A;Molecule type: DNA
A;Residues: 2-330 <HAR>
A;Cross-references: EMBL:Z17370
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of a
A;Reference number: S33887; MUID:83001943
A;Accession: S33887

A;Molecule type: DNA
A;Residues: 88-113;235-330 <TAK>
A;Cross-references: EMBL:Z17370
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman.
Biochemistry 9, 3161-3170, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A;Reference number: A90563; MUID:71064024
A;Contents: myeloma protein Eu
A;Accession: B90563

A;Molecule type: protein
A;Residues: 1-96,'R',98-135 <CUN>
A;Note: this sequence has the Gln(3) marker, 97-Arg
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se
A;Reference number: A90564; MUID:71064025
A;Contents: Eu
A;Accession: A90564
A;Molecule type: protein
A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2
A;Note: this sequence has the Gln(non-1) markers, 239-Glu and 241-Met
R;Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni
igen Primaerstruktur.
A;Reference number: A91668; MUID:77070269
A;Contents: myeloma protein Nie
A;Accession: B91668
A;Molecule type: protein
A;Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E'
A;Note: this sequence has the Gln(17) and Gln(1) markers
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A;Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1
A;Reference number: A91723; MUID:83289131
A;Contents: myeloma protein KOL; disulfide bonds
A;Accession: A91723
A;Molecule type: protein
A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <
A;Note: this sequence has the Gln(3) and Gln(non-1) markers
R;Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disul
A;Reference number: A90565; MUID:71064027
A;Contents: annotation; disulfide bonds
R;Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunog
enamide cleavage products, and the disulfide bridges.
A;Reference number: A91667; MUID:77070267
A;Contents: annotation; disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG1
A;Cross-references: GDB:120085; OMIM:147100
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 114/1; 224/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ha
in disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMI>
F;137-206/Domain: immunoglobulin homology <IM2>
F;243-310/Domain: immunoglobulin homology <IM3>
F;27-83,144-204,250-308/Disulfide bonds: #status experimental
F;103/Disulfide bonds: interchain (to light chain) #status experimental
F;109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F;180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 72.9%; Score 1224; DB 1; Length 330;
Best Local Similarity 97.4%; Pred. No. 5.4e-81;
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 80 RSCDKTHTCPCPAEAGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 139
Db :|||||
101 KSCDKTHTCPCPAELGPGSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNW 160
QY 140 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
Db :|||||
161 YVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 220
QY 200 KAKGPQREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPV 259

Db	221	KAGQPREPQYVLPSSRDELTKNQVSLTCLVGFYPSDIAVWESNGOPENNYKTPPV	280
Qy	260	LDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALNHHYTKQSLSPK	309
Db	281	LDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALNHHYTKQSLSPK	330

RESULT
A23511
4

Ig gamma-3 chain C region (allotype G3m(b)) - human
C/Species: Homo sapiens (man)
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C/Accession: A23511
R/Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A/Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: cd
A/Reference number: A23511; MUID:86148507
A/Accession: A23511
A/Molecule type: DNA
A/Residues: 1-377 <HUC>
A/Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056
C/Genetics:
A/Gene: GDB:IGHG3
A/Cross-references: GDB:119339; OMIM:147120
A/Map position: 14q32.33-14q32.33
A/Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin
F/20-85/Domain: immunoglobulin homology <IMM>

Query Match	69.8%;	Score 1171.5;	DB 2;	Length 377;
Best Local Similarity	83.0%;	Pred. No. 3.8e-77;		
Matches 220;	Conservative	9;	Mismatches 33;	Indels 3;
				Gaps 1;

Qy	48	KCMDASCRAHPHSDFC- --LCGAAAPPAPFRLLLWRSCKDHTPCPCPAPEAGAPSVEL	104
Dd	113	RCPKPKSCDTPPPCRPEPESCDTTPPCRCEPKSCTDPCCRCPACPELLGGPSVEL	172
Qy	105	FPPKPKOTLMSRTPEVTCVVVDVSHEDPEVKENWKYDVGEVHNATKKPREOYNSTFRV	164
Dd	173	FPPKPKOTLMSRTPEVTCVVVDVSHEDPEVQFKWIYDVGEVHNATKKPREOYNSTFRV	232
Qy	165	VSVLTVLHQDWLNKEYKCVKSNAKLPAPIEKTISKAKGPREPQVYTLPSPREEMTKNO	224
Dd	233	VSVLTVLHQDWLNKEYKCVKSNAKLPAPIEKTISKKGPREQPQVYTLPPSPREEMTKNO	292
Qy	225	VSLTCLVKGPYPDSIAVEWESNGPENNYKTTTPPLVSDSGSFFLYSKLTWDKSRWOQGNY	284
Dd	293	VSLTCLVKGPYPDSIAVEWESSQCPENNYNTTTPMLDSDGSFFLYSKLTWDKSRWOOGNI	352
Qy	285	FCSCVMHEALHNNHYTKLSLSLSPGK	309
Dd	353	FCSCVMHEALHNRFTOKLSLSLSPGK	377

RESULT 5

PT0207

Ig gamma chain C region - chimpanzee

C:Species: Pan troglodytes (chimpanzee)

C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999

C:Accession: PT0207

R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol.:Immunol. 28, 319-322, 1991

A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; MUID:91287716

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology <IMM>

Query Match	69.8%;	Score 1171;	DB 2;	Length 234;
Best Local Similarity	96.0%;	Pred. No. 2.4e-77;		
Matches 214;	Conservative	4;	Mismatches 5;	Indels 0; Gaps 0;

QY	80	RSCDKTHTCPPCPAPEAGAPSVLEFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNW	139
	:		
Db	12	KSCDTTHCTPCAPELLGGPSVELFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW	71
	:		
QY	140	YVDGVEVHNATKKPREEQYNSTYRVWSVLTVLHODWLNGKEYCKCVSNKALPAPIEKTIS	199
	:		
Db	72	YVDGVEVHNATKKPREEQYNSTYRVWSVLTVLHODWLNGKEYCKCVSNKALPAPIEKTIS	131
	:		
QY	200	KAKGQPPEQPYVTLPSPREEMTKNOVSLTCLVGKFYPSDIAVENESGNQPNNTKTPPV	259
	:		
Db	132	KAKGQPPEQPYVTLPSPSDELTKNQVSLTCLVGKFYPSDIAVENESSQPNNTKTPPV	191
	:		
QY	260	LDSGSEFFLYSKLVTDKSRWQQGVNFGSCVMHEALHHNYTKQS	302
	:		
Db	192	LDSGSEFFLYSKLVTDKSRWQQGVNFGSCVMHEALHHNYTKQS	234
	:		

RESULT 6

A60764
Ig gamma-3 chain C region, form LAT - human
C:Species: Homo sapiens (man)
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999
C:Accession: A60764
R:Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989
A:Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conv
A:Reference number: A60764; MUID:90007613
A:Accession: A60764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match	69.7%	Score 1169.5;	DB 2;	Length 377;
Best Local Similarity	83.0%	Pred. No. 5.3e-77;		
Matches 220;	Conservative	9;	Mismatches 33;	Indels 3;
				Gaps 1;

[illegible]

RESULT

```

G2HU
Ig gamma-2 chain C region : human
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1981 #sequence_revision 13-Jun-1983 #text_change 21-Jul-2000
C;Accession: A93906; A92809; A90752; A93132; A02148

```

R:Ellison, J.; Hood, L.
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con
 A:Reference number: A93906; MUID:82197621
 A:Accession: A93906
 A:Molecule type: DNA
 A:Residues: 1-326 <ELL>
 A:Cross-references: GB:V00554; GB:J00230; NID:932759; PIDN:CAB58438.1; PID:96066056
 A:Note: Lys-326 is probably removed posttranslationally
 R:Wang, A.C.; Tung, E.; Fudenberg, H.H.
 J. Immunol. 125, 1048-1054, 1980
 A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and
 A:Reference number: A92809; MUID:81007873
 A:Contents: myeloma protein T11
 A:Accession: A92809
 A:Molecule type: protein
 A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>
 A:Note: Trp-156 is at or near the complement-binding site
 R:Connell, G.E.; Parr, D.M.; Hofmann, T.
 Can. J. Biochem. 57, 758-767, 1979
 A:Title: The amino acid sequences of the three heavy chain constant region domains of a
 A:Reference number: A90752; MUID:80001357
 A:Contents: myeloma protein Z1e
 A:Accession: A90752
 A:Molecule type: protein
 A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, '222', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-
 A:Note: this sequence has since been revised
 R:Hofmann, T.; Parr, D.M.
 Mol. Immunol. 16, 923-925, 1979
 A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
 A:Reference number: A93132; MUID:80114419
 A:Contents: Z1e
 A:Accession: A93132
 A:Molecule type: protein
 A:Residues: 238-275 <HOF>
 R:Hofmann, T.; Parr, D.M.
 Submitted to the Atlas, March 1980
 A:Reference number: A94591
 A:Contents: annotation: Z1e, revisions to residues 25, 59, 60, and 264-268
 A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
 ned
 R:Milstein, C.; Frangione, B.
 Biochem. J. 121, 217-225, 1971
 A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.
 A:Reference number: A90253; MUID:72033500
 A:Contents: annotation: myeloma protein Sa, disulfide bonds
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.
 Nature 221, 145-148, 1969
 A:Title: Structural studies of immunoglobulin G.
 A:Reference number: A93157; MUID:69064124
 A:Contents: annotation; Sa, disulfide bonds
 C:Genetics:
 A:Gene: GDB:IGHG2
 A:Cross-references: GDB:119338; OMIM:147110
 A:Map position: 14q32.33-14q32.33
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:20-85/Domain: immunoglobulin homology <IMI>
 F:133-202/Domain: immunoglobulin homology <IM2>
 F:239-306/Domain: immunoglobulin homology <IM3>
 F:14/Disulfide bonds: Interchain (to light chain) #status experimental
 F:27-83, 140-200, 246-304/Disulfide bonds: #status experimental
 F:102, 103, 106, 109/Disulfide bonds: Interchain (to heavy chain) #status experimental
 F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.1%; Score 1159.5; DB 1; Length 326;
 Best Local Similarity 66.8%; Pred. No. 2.3e-76;
 Matches 229; Conservative 8; Mismatches 30; Indels 77; Gaps 5;
 QY 24 SVAGEQAPGTPAGSRGSSWSADLDKCMDCASCARPHSDFCLGCAAP--PAPFRLIWR 81

Db 2 STKGPSVFELAPCSRSTS-----ESTAAAGCLVKDYFPEPVTVSWNS 43
 QY 82 -----CDKTH----- 86
 Db 44 GALTSGVHTFPAVLQSSGLYSLSSVTVTPSSNFGTQYTCNVDPKPSNTKVDKTKVERKCC 103
 QY 87 -TCPGPCPAEAGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 145
 Db 104 VECPCPAPPVAG-PSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVE 162
 QY 146 VHNAKTKPREEQNSTYRVVSVLTVLDHQLNGKEYKCKVSNKALPAPIEKTTISKAKGQP 205
 Db 163 VHNAKTKPREEQNSTFRVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTTISKTKGQP 222
 QY 206 REQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTKTPPVLDSDGS 265
 Db 223 REQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTKTPPVLDSDGS 282
 QY 266 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 309
 Db 283 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 326
 RESULT 8
 G3HUWI
 Ig gamma-3 heavy chain disease proteins - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
 C:Accession: A90442; A92219; A90198; A93915; A02149
 R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
 Biochemistry 19, 4304-4308, 1980
 A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 hea
 A:Reference number: A90442; MUID:81021548
 A:Contents: heavy chain disease protein Wis
 A:Accession: A90442
 A:Molecule type: protein
 A:Residues: 1-289 <FRA>
 A:Note: the molecule is a dimer linked by 12 disulfide bonds: it has an extra interch
 A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12
 A:Note: the sequence of residues 42-76 was taken from the reference that follows
 R:Michaelsen, T.E.; Frangione, B.; Franklin, E.C.
 J. Biol. Chem. 252, 883-889, 1977
 A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplicat
 A:Reference number: A92219; MUID:77118561
 A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protei
 A:Accession: A92219
 A:Molecule type: protein
 A:Residues: 12-97 <MIC>
 A:Note: the hinge region in gamma-3 chains is about four times as long as in other ga
 idue segment (12-28)
 A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form in
 R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
 Biochem. Biophys. Res. Commun. 71, 907-914, 1976
 A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of t
 A:Reference number: A90198; MUID:77021516
 A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residu
 A:Accession: A90198
 A:Molecule type: protein
 A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>
 A:Note: this protein lacks most of the V region, all of the CH1 region, and part of t
 R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood,
 Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
 A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deleti
 A:Reference number: A93915; MUID:82247835
 A:Contents: heavy chain disease protein Omm
 A:Accession: A93915
 A:Molecule type: mRNA
 A:Residues: 12-70, 72-114, 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-
 A:Note: a carboxyl-terminal Lys is removed posttranslationally
 A:Note: this sequence may represent an allelic form or another gamma chain subclass
 C:Comment: The heavy chain disease protein Wis is shown.
 C:Genetics:

A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F:203-270/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 68.3%; Score 1146.5; DB 1; Length 289;
Best Local Similarity 81.1%; Pred. NO. 1.8e-75;
Matches 214; Conservative 14; Mismatches 33; Indels 3; Gaps 1;

QY 48 KCMDCASCRARPHSDFC---LGCAAAPAPAPFLLWRSCDKTHTCPCPCAPAPAGAPSVFL 104
DB 26 RCPEPKSCDTPPCPCPCPEPKSCDTPPCPCPEPKSCDTPPCPCPCAPAPFLLGSPSVFL 85
QY 105 FPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNAKTKPREQYNSTYRV 164
DB 86 FPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNAKTKPREQYNSTYRV 145
QY 165 VSVLTVLHODWLNKKEYCKVSNKALPAPIETKISKAKGQPREQVYTLPPSREEMTKNQ 224
DB 146 VSVLTVLHONLDGKEYCKVSNKALPAPIETKISKAKGQPREQVYTLPPSREEMTKNQ 205
QY 225 VSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 284
DB 206 VSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 265
QY 285 FSCSVMEALHNHYTQKSLSLSPG 308
DB 266 FSCSVMEALHNHYTQKSLSLSPG 289

RESULT 9
G4HU
Ig gamma-4 chain C region - human
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
C:Accession: A90933; A90249; A02150
R:Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A:Reference number: A90933; MUID:83157104
A:Accession: A90933
A:Molecule type: DNA
A:Residues: 1-327 <ELL>
A:Note: the sequence was determined from the germline gene
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A:Title: Human immunoglobulin subclases. Partial amino acid sequence of the constant r
A:Reference number: A90249; MUID:70207560
A:Accession: A90249
A:Molecule type: protein
A:Residues: 1-30;81-326 <PIN>
C:Genetics:
A:Gene: GDB:IGHG4
A:Cross-references: GDB:119340; OMIM:147130
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 111/1; 221/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 14
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:99-110/Region: hinge
F:134-203/Domain: immunoglobulin homology <IM2>
F:240-307/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (co light chain) #status experimental
F:27-83,141-201,247-305/Disulfide bonds: #status predicted
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.9%; Score 1139; DB 1; Length 327;
Best Local Similarity 65.7%; Pred. NO. 7e-75;
Matches 226; Conservative 9; Mismatches 33; Indels 76; Gaps 4;

QY 24 SVAGQAPGTCAPCSRGSWSADLDCMDCASCRARPHSDFCGLCAAAP--PAPFRLWRS 81
DB 2 STKGFSVPFLAPCSRSTS-----ESTALGCLVKDYFPEPVTWSNS 43
QY 82 -----CDKTH----- 86
DB 44 GALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTTTCNVNDHKPSNTKVKRVESKYG 103
QY 87 -TCPPCPAPAEAGAPSVFLPPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGYE 145
DB 104 PPCPCAPAEFLGGPSVFLPPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGYE 163
QY 146 VHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETKISKAKGP 205
DB 164 VHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKGLPSSIEKTIISKAKGP 223
QY 206 REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 265
DB 224 REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 283
QY 266 PFLYSLKLVDRKRWQGNVFCFSVMHEALHNHYTQKSLSLSPG 309
DB 284 PFLYSLKLVDRKRWQGNVFCFSVMHEALHNHYTQKSLSLSPG 327

RESULT 10
147159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845
A:Accession: I47159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:9433123; PIDN:AAA52217.1; PID:9433124
C:Genetics:
A:Gene: IgG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 55.3%; Score 928; DB 2; Length 328;
Best Local Similarity 52.7%; Pred. NO. 1.1e-59;
Matches 183; Conservative 36; Mismatches 46; Indels 82; Gaps 7;

QY 26 AGEQAPGT---APCSRGSWSADLDCMDCASCRARPHSDFCGLCAAAP--PAPFRLWLR 80
DB 1 APKTAHSVYPLAPCSRSTS-----GPNVALGCLASSYFPEPVTVTWN 42
QY 81 S-----CDKTH----- 86
DB 43 SGALSSGVHTFPAVLQSSGLYSLSSVTVTPASSLSKSYTCNVNHPATTTKVKDKRGVTKT 102
QY 87 --TCPPCPAPAEAGAPSVFLPPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGY 144
DB 103 KPPCPICPACESPG-PSVFIFPPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGY 161
QY 145 EVHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETKISKAKQ 204
DB 162 EVHTAQTRKPEQFNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETKISKAKQ 221
QY 205 PREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQ--PENNYKTTPPVLD 262

```

Db      222 TRPEQYTLPPHAEELSRKSVITCLVIGFYPPDIDVEWRNGQPPEEGNYRTTTPQQDV 281
           ||||| | | : : : : : ||| | | | | | | | | | | | | | | | | | | | | |
Qy      263 DGSFFLYSLTVDKSRWQGNSFCVSVMHEALHNHYTKQSLSLSPGK 309
           ||||| | | | : | | | | | | | | | | | | | | | | | | | | | |
Db      282 DGTYFLYSFSDVKASWGQGGIFQCAGMHEALHNHYTKQSISKTPGK 328
           ||||| | | | : | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
I47160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47160
R:Kacs Kovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:title: Five putative subclasses of swine IgG identified from the CDNA sequence
A:Reference number: I47158; MUID:95015845
A:Accession: I47160
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:g433126; PIDN:RAAS2218.1; PID:g433126
C:Genetics:
A:Gene: IgG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:I33-202/Domain: immunoglobulin homology <IMX>

Query Match          55.0%; Score 923; DB 2; Length 328;
Best Local Similarity 52.4%; Pred. No. 2.5e-59;
Matches 182; Conservative 35; Mismatches 44; Indels 86; Gaps

Qy      30 APGT-----APCSRGSSWSADLDKCMDCASCRRPHSDFCLGCAAAP--PAPPELLWR 80
           || | | | | | : | | | | | : | | | | | | | | | | |
Db      1 APKTAPLVPLAPCGRDTS-----GNVALGCLASSIFYPEPVVTWN 42
           || | | | | | : | | | | | : | | | | | | | | | | |
Qy      81 S-----CDKTH----- 86
           | | | | | | | | | | | | | | | | | | | | | |
Db      43 SGALTSCVHTFPVLPSGLYLSSMVTPASSLSKSVCNVNHPTTKVDKRVGTKT 102
           | | | | | | | | | | | | | | | | | | | | | |
Qy      87 --TCPCPAEAGAPSFVFPPPKPDLTMSIPTCVTVVDSVEDPEVKFNMYVDGV 144
           || || | | | | | | | | | | | | | | | | | | | | |
Db      103 KPCCPICACESPG-PSVFIPFKPKDRLTMSIRTPQVTCVVVDVSGENPEVQFSWYVDGV 161
           || || | | | | | | | | | | | | | | | | | | | | |
Qy      145 EVHNATKPREEQNSTYRVSVVLTVLHQDLNGKEYCKKCVSNKALPAPIETIKSKAQG 204
           || | | | | | | | | | | | | | | | | | | | | | |
Db      162 EVAHTAOTRPKEGFNSTYRVSVVLPHQDLNGKEFKCKVNKDLPAPIRIISKAQG 221
           || | | | | | | | | | | | | | | | | | | | | | |
Qy      205 PREPOVYTLPSPREMTKNQSVLCLKVGYPSDIADVESNQ--PENNYKTTTPVLDS 262
           ||||| | | | | | | | | | | | | | | | | | | | | |
Db      222 TREPOVYTLPHAEELSRKSVITCLVIGFYPPDIDVEWRNGQPPEEGNYRTTTPQQDV 281
           ||||| | | | : | | | | | | | | | | | | | | | | |

Qy      263 DGSFFLYSLTVDKSRWQGNSFCVSVMHEALHNHYTKQSLSLSPGK 309
           ||||| | | | : | | | | | | | | | | | | | | | | |
Db      282 DGTYFLYSFSDVKASWGQGGIFQCAGMHEALHNHYTKQSISKTPGK 328
           ||||| | | | : | | | | | | | | | | | | | | | | |

RESULT 12
GBRB
Ig gamma chain C region - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A:title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
A:Reference number: A91749; MUID:84030930
A:Accession: A91749
A:Molecule type: mRNA
A:Residues: 1-323 <BER>
A>Note: This sequence has the d12 allotypic marker, 104-Thr, and the e14 meta
R:Pratt, D.M.; Mole, L.E.
```

[illegible]

C:Accession: I47162
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845
A:Accession: I47162
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <KAC>
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130
C:Genetics:
C:Gene: IgG4
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 54.1%; Score 907; DB 2; Length 277;
Best Local Similarity 72.1%; Pred. No. 2.9e-58;
Matches 165; Conservative 29; Mismatches 31; Indels 4; Gaps 3;

QY 85 HTCPCP-APEAEG-APSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYVD 142
DB 49 TKTRPCPCACEGPGSAFFPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYVD 108
QY 143 GVEVHNAKTPREQYNTYRVSVLTVLHODWLNKGEKCKVSNKALPAPIETISKAK 202
DB 109 GVEVHNAKTPREQYNTYRVSVLTVLHODWLNKGEKCKVSNKALPAPIETISKAK 168
QY 203 GPREQVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPPV 260
DB 169 GQTRPQVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPPV 228
QY 261 DSDGFELYSLTVDKSRWQGNVFCSCVMHEALHNYTQKSLSPGK 309
DB 229 DSDGFELYSLTVDKSRWQGNVFCSCVMHEALHNYTQKSLSPGK 277

RESULT 14
I47158
Ig gamma 1 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
A:Reference number: I47158
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845
A:Accession: I47158
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03778; NID:9433121; PIDN:AAA52216.1; PID:9433122
C:Genetics:
C:Gene: IgG1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 53.3%; Score 895; DB 2; Length 328;
Best Local Similarity 52.2%; Pred. No. 2.6e-57;
Matches 181; Conservative 31; Mismatches 53; Indels 82; Gaps 7;

QY 26 AGEQAPGT---APCSRGSSWSADLCKMDCASCARPHSDFCGLCAAAP--PAPFRLLR 80
DB 1 APKTAPSVYPLAPCGRDVS-----GPNVALGCLASSYFPEPVTMTWN 42
QY 81 S-----CDKTH----- 86
DB 43 SCALTSVHTPSPVLPQSLGLYSLSNMVTPASSLSKSYTCNNHPATTKVDRKRGVHQ 102

QY 87 --TCPPCPAPEAEGAPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYVDG 144
DB 103 PQTCPICPCGCEVAG--PSVFIAPPKPKDMLISQTPETVTCVVVDVSKHEAEVQFSWYVDG 161

QY 145 EVHNAKTPREQYNTYRVSVLTVLHODWLNKGEKCKVSNKALPAPIETISKAKQ 204
DB 162 EVHTAETRPKEQFNTYRVSVLTVLHODWLNKGEKCKVSNKALPAPIETISKAIQ 221
QY 205 PREPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPPV 262
DB 222 SREPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPPV 281
QY 263 DGSFFELYSLTVDKSRWQGNVFCSCVMHEALHNYTQKSLSPGK 309
DB 282 DGSFFELYSLTVDKSRWQGNVFCSCVMHEALHNYTQKSLSPGK 328

RESULT 15
I47161
Ig gamma 3 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
A:Accession: I47161
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845
A:Accession: I47161
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03781; NID:9433127; PIDN:AAA52219.1; PID:9433128
C:Genetics:
C:Gene: IgG3
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 53.2%; Score 893; DB 2; Length 328;
Best Local Similarity 52.3%; Pred. No. 3.6e-57;
Matches 181; Conservative 33; Mismatches 52; Indels 80; Gaps 8;

QY 26 AGEQAPGT---APCSRGSSWSADLCKMDCASCARPHSDFCGLCAAAP--PAPFRLLR 80
DB 1 APKTAPSVYPLAPCGRDVS-----GPNVALGCLASSYFPEPVTMTWN 42
QY 81 S-----CDKTH----- 88
DB 43 SCALTSVHTPSPVLPQSLGLYSLSNMVTPASSLSKSYTCNNHPATTKVDRKRGVTK 102

QY 89 --PPCP-APEAEG-APSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 145
DB 103 KPPCPICPCGCEVAGPSVFIAPPKPKDMLISQTPETVTCVVVDVSKHEAEVQFSWYVDGVE 162
QY 146 VHNATKPREQYNTYRVSVLTVLHODWLNKGEKCKVSNKALPAPIETISKAKQ 205
DB 163 VHTAETRPKEQFNTYRVSVLTVLHODWLNKGEKCKVSNKALPAPIETISKAIGS 222
QY 206 REPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPPV 263
DB 223 REPOVYTLPPSREEMTKNOVSLTCLVKGFPYSDIAVESNGQ--PENNYKTTTPPV 282

QY 264 SFFELYSLTVDKSRWQGNVFCSCVMHEALHNYTQKSLSPGK 309
DB 283 GTFFELYSLTVDKSRWQGNVFCSCVMHEALHNYTQKSLSPGK 328

Search completed: March 11, 2002, 15:14:51
Job time: 135 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:12:36 ; Search time 53.92 Seconds
(without alignments)
424,493 Million cell updates/sec

Title: US-09-742-454A-7

Perfect score: 1678

Sequence: 1 MARGSLRLRLVLGLWLA.....MHEALHHYTKSLSLSPGK 309

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1678	100.0	309	AAU03500	Human TWEAK receptor
2	1258	75.0	451	AA14765	NAIL-Fc protein se
3	1255	74.8	438	AA181882	Plasmid pDC406/Ox4
4	1255	74.8	438	AA181882	Ox40/Fc mutin. C
5	1250	74.5	660	AA113463	Ectromelia A39R se
6	1250	74.5	660	AA113463	Ectromelia A39R se
7	1250	74.5	660	AA113463	Ectromelia A39R se
8	1250	74.5	660	AA113463	Ectromelia A39R se
9	1247	74.3	376	AA160037	Amino acid sequenc
10	1247	74.3	450	AA15239	Antigenic peptide
11	1247	74.3	453	AA15240	ULBP1-Fc fusion po

12	1243	74.1	424	16	AAW14764	Human soluble kit
13	1243	74.1	424	16	AAW14765	Human soluble kit
14	1239.5	73.9	375	18	AAW06683	Heregulin-alpha fu
15	1239	73.8	488	21	AAW97175	Human FGE-R1 Extra
16	1239	73.8	535	20	AAW17414	SYPH1-26 disintegr
17	1235.5	73.6	234	18	AAW16686	Human IgG1 Fc port
18	1235	73.6	400	21	AAW15123	Porcine CTLA-4-Ig
19	1233.5	73.5	396	18	AAW18574	Aggrecanase artifi
20	1233.5	73.5	396	18	AAW18575	Aggrecanase artifi
21	1232	73.4	376	18	AAW050287	Human Fas antigen
22	1230.5	73.3	302	22	AAE00507	Human BCMA-Immunog
23	1230.5	73.3	302	22	AAE00507	Mouse IgG signal/h
24	1230.5	73.3	302	22	AAE00507	Human Interleukin
25	1230.5	73.3	302	22	AAE00507	Human Interleukin
26	1230	73.3	292	22	AAE02642	Human immunoglobul
27	1230	73.3	292	22	AAE02642	Human immunoglobul
28	1230	73.3	292	22	AAE02642	Human IgG1 Fc regi
29	1230	73.3	330	20	AAW0153	Human heavy chain
30	1230	73.3	339	22	AAE02648	Human immunoglobul
31	1230	73.3	401	18	AAW10537	Leptin 1-167/IgG1
32	1230	73.3	444	21	AAW32263	Humanised anti-CD2
33	1230	73.3	451	20	AAW50031	Human E27 anti-IgE
34	1230	73.3	451	20	AAW50031	Mus musculus anti-
35	1230	73.3	451	20	AAW50031	Mus musculus anti-
36	1230	73.3	451	20	AAW50031	Mus musculus anti-
37	1230	73.3	451	21	AAW85201	Amino acid sequenc
38	1230	73.3	451	21	AAW85201	Light chain amino
39	1230	73.3	451	22	AAW74212	E27 anti-IgE antib
40	1230	73.3	451	22	AAW74212	Anti-IgE antibody,
41	1230	73.3	451	22	AAW76948	Full length heavy
42	1230	73.3	451	22	AAW76950	Full length heavy
43	1230	73.3	451	22	AAW76952	Full length heavy
44	1230	73.3	452	19	AAW69316	Anti-IL-8 humanise
45	1230	73.3	452	20	AAW29458	Recombinant immuno
				21	AAW30322	Humanised anti-IL-

ALIGNMENTS

RESULT 1

AAU03500
ID AAU03500 standard; Protein; 309 AA.

AC AAU03500;

XX 26-SEP-2001 (first entry)

XX Human TWEAK receptor-Fc (TWEAKR-Fc) fusion polypeptide.

DE TWEAK receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;
KW corneal graft neovascularisation; psoriasis; metastatic condition;
KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
KW preneoplastic condition; myocardial angiogenesis; wound granulation;
KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;
KW peripheral atherosclerosis; human IgG1; TWEAKR-Fc; fusion protein.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers
FT Peptide 1..27 /note= "Signal peptide"

FT Protein 28..309 /note= "Mature human TWEAKR-Fc fusion protein."

FT Domain 28..79 /note= "Specifically referred to in Claim 11"

FT Region 80..81
/note= "From TWEAKR extracellular domain. Specifically referred to in Claims 4 and 10"


```

Query Match      75.08; Score 1258; DB 20; Length 451;
Best Local Similarity 86.3%; Pred. No. 1:2e-87;
Matches 240; Conservative 3; Mismatches 15; Indels 20; Gaps 3;

QY 33 TAPSRGSSW-SADLDKCMDCASCRRARPHSDFCGLGCAAAPAPAPFRLLRSCDKTHTCPPC 91
DB 193 tcnvsnpsweshlnltqdcqna---hqef-----rrscdkthtcppc 233
QY 92 PAPAEGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKT 151
DB 234 papaegapsvflfppkpkdtlmisrtpevtcvvdvshedpevkfnwvvgvevhnaht 293
QY 152 KPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVY 211
DB 294 kpreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektiskakgqprepqvy 353
QY 212 TLPPSREEMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK 271
DB 354 tlppsreemtknqvsltclvkgfypsdiavewesngqpennnykttppvldsdgsfflysk 413
QY 272 LTVKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 309
DB 414 ltvksrwqqgnvfscsvmhealhnhytqkslsispgk 451

RESULT 3
AAR81882
ID AAR81882 standard; Protein; 438 AA.
XX AC AAR81882;
XX DT 30-MAR-1996 (first entry)
XX DE Plasmid pDC406/OX40/Fc* encoding an OX40/Fc mutein protein.
XX KW OX40; OX40-L; cytokine; cell surface molecule; plasmid;
XX KW pDC406/OX40/Fc*; membrane glycoprotein.
XX OS Synthetic.
XX PN US5457035-A.
XX PD 10-OCT-1995.
XX PF 23-JUL-1993; 93US-0097827.
XX PR 23-JUL-1993; 93US-0097827.
XX PA (IMMV ) IMMUNEX CORP.
XX PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
XX DR WPI; 1995-357992/46.
XX DR N-PSDB; AAT00829.
XX PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
XX PT and host cells, used to produce recombinant ligand used in e.g.
XX PT prim. T cell culture, to modulate immune response etc.
XX PS Example 2; Column 35-38; 26pp; English.
XX CC This plasmid encodes an OX40/Fc antibody fragment mutein protein,
XX CC and is used to express a soluble OX40/Fc mutein fusion protein for
XX CC use in detecting cDNA clones encoding a OX40 ligand. The Fc
XX CC fragment may be derived from human IgG1, and the plasmid may be
XX CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
XX CC line. Culture supernatant was purified by affinity chromatography
XX CC and this was used, together with labeled goat anti-human IgG to
XX CC screen various cell lines.
XX SQ Sequence 438 AA;

```

```

Query Match      74.8%; Score 1255; DB 16; Length 438;
Best Local Similarity 65.8%; Pred. No. 2e-87;
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;

QY 35 PCSRG-SSWSADLDKCMDCAS-----CRARP-----HSDFCLG- 66
DB 62 pcegtfyneavnydckqctqcnhrsgselkqncptqtdtvcrcrpgtqqrqdsykgiv 121
QY 67 -CAAAPAPF-----RLLM----- 79
DB 122 dcvcpcpgfhfsgpnqackpwtnctlsqkqtrhpasdsldavcedrsllatillwetqrpt 181
QY 80 -----RSCDKTHTCPPCPAPEAGAPSVFLFPPKPKDT 112
DB 182 frptvtqgttvpwrtseipstplveprscdkthccppcpapeagapsvflfppkpkdt 241
QY 113 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTTPREQYNSTYRVVSVLTVLH 172
DB 242 lmisrtpevtcvvdvshedpevkfnwvvgvevhnahtkpreeqynstyrvvsvltvlh 301
QY 173 QDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVK 232
DB 302 qdwlngkeyckvsnkalpapiektiskakgqprepqvytlppsrdeltknqsvltclvk 361
QY 233 GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHE 292
DB 362 gfypsdiavewesngqpennnykttppvldsdgsfflyskltdvksrwqqgnvfscsvmhe 421
QY 293 ALHNHYTQKSLSLSPGK 309
DB 422 alhnhytqkslsispgk 438

RESULT 4
AAW48976
ID AAW48976 standard; Protein; 438 AA.
XX AC AAW48976;
XX DT 25-SEP-1998 (first entry)
XX DE OX40/Fc mutein.
XX KW OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;
XX KW chimeric.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Mus sp.
XX FH Key Location/Qualifiers
XX FT Region 1..206
XX FT /note= "Extracellular domain of mouse OX40"
XX FT Region 207..438
XX FT /note= "Mutant Fc region of human IgG1 antibody"
XX FT Misc-difference 225
XX FT /note= "changed from Leu in wild-type to Ala in
XX FT mutant"
XX FT Misc-difference 226
XX FT /note= "changed from Leu in wild-type to Gly in
XX FT mutant"
XX FT Misc-difference 228
XX FT /note= "changed from Gly in wild-type to Ala in
XX FT mutant"
XX PN US5783665-A.
XX PD 21-JUL-1998.
XX PF 22-JUN-1995; 95US-0494574.
XX PR 23-JUL-1993; 93US-0097827.
XX PR 22-JUN-1995; 95US-0494574.

```

```
XX PA (IMMV ) IMMUNEX CORP.
XX PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
XX DR WPI; 1998-427099/36.
XX DR N-PSDB; AAV32636.
XX PT Purified polypeptide OX-40 ligands - for co-stimulation of T-cell
XX PT production and binding assays for OX-40 and homologues
XX PS Example 2; Col 37-40; 26pp; English.
XX CC The present sequence represents the OX40/Fc fusion protein that
XX CC contains the extracellular domain of mouse OX40 fused to the mutated
XX CC Fc region of the human IgG1 antibody. The fusion protein was used
XX CC for detecting cDNA clones encoding an OX40 ligand. The invention
XX CC claims for a murine OX40-L cytokine (AAW48975) that binds to the murine
XX CC T cell antigen, OX40. The OX40-L protein is claimed to be useful for
XX CC co-stimulation of T-cell production and in binding assays for
XX CC detecting OX40 or its homologues. The OX40-L protein is also claimed
XX CC to generate a TH-2 immune response.
XX SQ Sequence 438 AA;
Query Match 74.8%; Score 1255; DB 19; Length 438;
Best Local Similarity 65.8%; Pred. No. 2e-87;
Matches 248; Conservative 6; Mismatches 21; Indels 102; Gaps 6;
QY 35 PCSRG-SSWADLDKCMDCAS-----CRAP-----HSDFCLG- 66
DB 62 pcegtfyneavnydtkqctqcnhrsgselkqncptqdtvcrpqtqprqdsykglv 121
QY 67 -CAAPAPF-----RLW----- 79
DB 122 dvcpcppghspgnnqackpwtncitlsgkqtrhpasdsldavcedrslallilwetqrpt 181
QY 80 -----RSCDKHTHTCPPCPAPEGAPSVFLPPPKPDT 112
DB 182 frptvgsttwrtseipstptlveprscdkthtccpapeaegapsvflppkpkdt 241
QY 113 LMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVEVHNNAKTRPEEQYNSTYRVVSVLTVL 172
DB 242 lmsrtpevtcvcvvdshedpevkfnwydgvvevhnaktpreeqynstyrsvsvltvlh 301
QY 173 ODWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDEMTKNQVSLTCLVK 232
DB 302 qdwlngkeykckvsnkalpapiektiskakgqprepqvyltppsrdeitknqsvltclvk 361
QY 233 GFYPDSIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFSV 292
DB 362 gfypdsiavewesngqpennnykttppvldsdgsfflyskltvdksrwqqgnvfscvmhe 421
QY 293 ALHNHYTQKSLSLSPGK 309
DB 422 alhnhytqkslsispkg 438
RESULT 5
AAV13463
ID AAV13463 standard; Protein; 660 AA.
XX AC AAV13463;
XX DT 26-JUL-1999 (first entry)
XX CC Ectromelia A39R semaphorin polypeptide.
XX KW VESPR; viral-encoded semaphorin protein receptor; semaphorin; IL-12;
XX KW inflammatory disease; viral infection; immune regulation; interleukin-12;
XX KW Th1 type immune response; cancer; autoimmune disease; A39.
```

```
OS Ectromelia virus.
XX PN WO9921997-A1.
XX PD 06-MAY-1999.
XX PF 28-OCT-1998; 98WO-US22879.
XX PR 28-OCT-1997; 97US-0112009.
XX PA (IMMV ) IMMUNEX CORP.
XX PI Comeau MR, Dubose RE, Johnson RS, Spriggs MK;
XX DR WPI: 1999-326704/27.
XX DR N-PSDB; AAX55565.
XX PT Viral-encoded protein receptors useful for treating inflammatory
XX PT diseases
XX PS Example 1; Page 63-66; 73pp; English.
XX CC The invention relates to a VESPR (viral-encoded semaphorin protein
XX CC receptor) polypeptide that can bind semaphorins. The VESPR polypeptides
XX CC are used to treat inflammatory diseases; when immobilized, to purify and
XX CC separate semaphorins or cells that express them. They are useful as
XX CC reagents for detecting, or measuring biological activity of, optionally
XX CC modified semaphorins; as carriers for delivering diagnostic and
XX CC therapeutic agents to semaphorin-positive cells; for studying the role of
XX CC VESPR in viral infection and immune regulation. When administered, in
XX CC combination with vaccination, together with a semaphorin, VESPR may
XX CC stimulate production of interleukin-12 (IL-12) which acts as an adjuvant
XX CC to induce a more persistent cellular immune response (of Th1 type),
XX CC including a curative response against aggressive, micrometastatic
XX CC cancers. Antibodies against VESPR can be used to treat autoimmune
XX CC diseases where an inflammatory response follows presentation of self-
XX CC antigens by T cells. Sense and antisense fragments of nucleic acid
XX CC encoding VESPR can be used to block expression of the polypeptide. The
XX CC present sequence represents an eromelia A39 semaphorin protein. The
XX CC native VESPR polypeptide is isolated from the human cells expressing the
XX CC receptor by using an Ectromelia virus A39 semaphorin/fc Fusion protein.
XX SQ Sequence 660 AA;
Query Match 74.5%; Score 1250; DB 20; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.8e-87;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 80 RSCDKHTHTCPPCPAPEGAPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNW 139
DB 34 rscdkthtccpapeaegapsvflppkpkdtlmsrtpevtcvcvvdshedpevkfnw 93
QY 140 YDGVGEVHNNAKTRPEEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS 199
DB 94 ydgvgevhnnaktpreeqynstyrsvsvltvlhqdwlngkeykckvsnkalpapiektis 153
QY 200 KAKGQPREPOVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 259
DB 154 kagqprepqvyltppsrdeemtknqsvltclvkgyfypsdlavewesngqpennnykttppv 213
QY 260 LDSGSGFFLYSLKLTVDKSRWQQGNVFSVSCSYMHEALHNHYTQKSLSLSPGK 309
DB 214 ldsdgsfflyskltvdksrwqqgnvfscsvmhealhnhytqkslsispkg 263
RESULT 6
AAB28523
ID AAB28523 standard; Protein; 660 AA.
XX AC AAB28523;
XX DT 07-FEB-2001 (first entry)
```

XX DE Ectromelia A39R semaphorin.
 XX DE Ectromelia; semaphorin; VESPR; viral encoded semaphorin protein receptor;
 KW anti-inflammatory; semaphorin modulator; rheumatoid arthritis;
 KW inflammation; immune regulation; viral infection.
 XX OS Ectromelia virus.
 XX PN US6130068-A.
 XX PD 10-OCT-2000.
 XX PF 28-OCT-1998; 98US-0181706.
 XX PR 26-OCT-1998; 98US-0112009.
 XX PA (IMMV) IMMUNEX CORP.
 XX PI Comeau MR, Johnson RS, Spriggs MK, Dubose RF;
 XX WPI: 2000-646753/62.
 XX DR N-PSDB; AAC63726.
 XX PT New Viral Encoded Semaphorin Protein Receptor DNA and polypeptides,
 PT useful for treating inflammation or inflammatory diseases, e.g.
 PT rheumatoid arthritis -
 XX Example 1; Column 55-60; 32pp; English.
 XX The present sequence is provided in an invention relating to novel
 CC semaphorin receptor polypeptides designated Viral Encoded Semaphorin
 CC Protein Receptor (VESPR). The DNA encoding the VESPR is useful for
 CC constructing expression vectors that produce the VESPR polypeptide. The
 CC VESPR DNA and polypeptide are useful for treating inflammation and
 CC inflammatory diseases, e.g. rheumatoid arthritis, a disease associated
 CC with chronic inflammation of the synovial tissue. The VESPR polypeptide
 CC is useful for measuring the biological activity of semaphorin proteins
 CC in terms of their binding affinity for VESPR. It is also useful as a
 CC research tool for studying the role that the receptor, in conjunction
 CC with semaphorins, may play in immune regulation and viral infection. The
 CC VESPR polypeptides are also useful as reagents that may be employed by
 CC those conducting quality assurance studies, e.g. to monitor shelf life
 CC and stability of semaphorin protein under different conditions.
 XX SQ Sequence 660 AA;

Query Match 74.5%; Score 1250; DB 21; Length 660;
 Best Local Similarity 100.0%; Pred. No. 7.8e-87;
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 RSCDKTHTCCPCPAPEAGAPSVFLFPKPKDTLMISRPEVTCVVVDVSHEDPEVKFNW 139
 DB 34 rscdkthtccpcpapeagapsvflfpkpkdtlmsrtpetvctcvvdvshedpevkfnw 93
 QY 140 YVDGVEVHNKTPREEQNSTYRVVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTIS 199
 DB 94 yvdgvevhnaktpreeqnystyrvvsvltvlhqdwlngkeyckvsnkalpapiektis 153
 QY 200 KAKGQPREQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 259
 DB 154 kkgqpreqvtytlppsreemtknqvs lclvkgfy psdiavwesngopennykttppv 213
 QY 260 LQSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 309
 DB 214 lqsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispkg 263

RESULT 7
 AAB70132
 ID AAB70132 standard; Protein; 660 AA.
 XX

AC AAB70132;
 XX 21-MAY-2001 (first entry)
 DE Ectromelia A39R semaphorin.
 XX Ectromelia; semaphorin; VESPR; viral encoded semaphorin protein receptor;
 KW anti-inflammatory; semaphorin modulator; rheumatoid arthritis;
 KW inflammation; immune regulation; viral infection.
 XX OS Ectromelia virus.
 XX PN US6187909-B1.
 XX PD 13-FEB-2001.
 XX PF 10-DEC-1999; 99US-0459066.
 XX PR 26-OCT-1998; 98US-0112009.
 XX PR 28-OCT-1998; 98US-0181706.
 XX PR 28-OCT-1997; 97US-0958598.
 XX PA (IMMV) IMMUNEX CORP.
 XX PI Spriggs MK, Comeau MR, Dubose RF, Johnson RS;
 XX WPI: 2001-256141/26.
 XX DR N-PSDB; AAF76950.
 XX PT New viral encoded semaphorin protein receptor polypeptides useful for
 PT treating inflammation or diseases associated with pro-inflammatory
 PT activity of a semaphorin ligand -
 XX Example 1; Column 55-60; 32pp; English.
 XX The DNA encoding the present sequence was used to prepare an Ectromelia
 CC semaphorin/Fc fusion protein. The fusion protein was used in the
 CC isolation of a human viral encoded semaphorin protein receptor (VESPR).
 CC The VESPR polypeptides are useful for treating inflammation, or diseases
 CC associated with pro-inflammatory activity of a semaphorin ligand. The
 CC polypeptides are also useful as carriers for delivering diagnostic or
 CC therapeutic agents to cells expressing semaphorins.
 XX SQ Sequence 660 AA;

Query Match 74.5%; Score 1250; DB 22; Length 660;
 Best Local Similarity 100.0%; Pred. No. 7.8e-87;
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 RSCDKTHTCCPCPAPEAGAPSVFLFPKPKDTLMISRPEVTCVVVDVSHEDPEVKFNW 139
 DB 34 rscdkthtccpcpapeagapsvflfpkpkdtlmsrtpetvctcvvdvshedpevkfnw 93
 QY 140 YVDGVEVHNKTPREEQNSTYRVVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTIS 199
 DB 94 yvdgvevhnaktpreeqnystyrvvsvltvlhqdwlngkeyckvsnkalpapiektis 153
 QY 200 KAKGQPREQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPV 259
 DB 154 kkgqpreqvtytlppsreemtknqvs lclvkgfy psdiavwesngopennykttppv 213
 QY 260 LQSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 309
 DB 214 lqsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispkg 263

RESULT 8
 AAB31694
 ID AAB31694 standard; Protein; 660 AA.
 XX
 AC AAB31694;
 XX

DT 30-APR-2001 (first entry)
 XX Amino acid sequence of the Ectromelia virus A39R semaphorin.
 DE
 XX
 XX Viral encoded semaphorin protein receptor; VESPR; semaphorin; interferon;
 KW viral semaphorin ligand; Ectromelia A39R; Staphylococcus aureus; IL-12;
 KW interleukin-12; dendritic cell; cytokine production; immunomodulator;
 KW proinflammatory cytokine; Th1 cell differentiation; inflammation;
 KW inflammatory disease; tumour.
 XX
 XX Ectromelia virus.
 OS
 XX
 XX US6174689-B1.
 PN
 XX
 XX 16-JAN-2001.
 PD
 XX
 XX 10-DEC-1999; 99US-0458791.
 XX
 XX 26-OCT-1998; 98US-0112009.
 PR
 XX 28-OCT-1998; 98US-0181706.
 PR
 XX 27-OCT-1997; 97US-0958598.
 PR
 XX (IMV) IMMUNEX CORP.
 PA
 XX
 XX Spriggs MK, Comeau MR, Dubose RF, Johnson RS;
 PI
 XX
 XX WPI; 2001-167789/17.
 DR
 XX N-PSDB; AAF25254.
 DR
 XX
 XX Screening for binding to viral encoded semaphorin protein receptor,
 PT comprises contacting a mixture containing semaphorin, or cells that
 PT express semaphorin with protein, and detecting binding to protein
 PT
 XX
 XX Example 1; Column 55-60; 32pp; English.
 PS
 XX
 XX The present sequenc represents an Ectromelia virus A39R semaphorin
 CC ligand. The protein was used to isolate a viral encoded semaphorin
 CC protein receptor (VESPR) polypeptide from human cells. Semaphorins
 CC interact with their membrane bound receptors to synergise with interferon
 CC and Staphylococcus aureus (type C) in the production of interleukin-12
 CC (IL-12) from dendritic cells. VESPR can therefore be used to induce
 CC IL-12 production which in turn promotes cytokine production. IL-12 is
 CC a proinflammatory cytokine and an immunomodulator. A soluble VESPR can
 CC be used to antagonise IL-12 and downregulate Th1 cell differentiation.
 CC VESPR is therefore useful for treating inflammation and inflammatory
 CC diseases. VESPR ligands may be used to induce an immune response against
 CC aggressive tumours.
 CC
 XX
 XX Sequence 660 AA;

Query Match 74.5%; Score 1250; DB 22; Length 660;
 Best Local Similarity 100.0%; Pred. No. 7.8e-87;
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 80 RSCDKTHTCCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 139
 DB 34 RSCDKTHTCCPAPAEAGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 93
 QY 140 YVDGVEVHNKTRPREQYNSTYRVSVLTVLHODWLNKREYCKVSNKALPAPIEKTIS 199
 DB 94 YVDGVEVHNKTRPREQYNSTYRVSVLTVLHODWLNKREYCKVSNKALPAPIEKTIS 153
 QY 200 KAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 259
 DB 154 KAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 213
 QY 260 LDSGSEFFLYSKLTVDKSRWQQGNVFCSMVHEALHNHYTQKSLSLSPGK 309
 DB 214 LDSGSEFFLYSKLTVDKSRWQQGNVFCSMVHEALHNHYTQKSLSLSPGK 263

RESULT 9

AAW60037
 ID AAW60037 standard; Protein; 376 AA.
 XX
 AC AAW60037;
 XX
 DT 11-SEP-1998 (first entry)
 XX
 XX Antigenic peptide hFas (nd29) containing Fc region.
 DE
 XX
 KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
 KW heart failure; kidney failure; graft-versus-host disease; antibody;
 KW myocardial infarction; ischemic restenosis; endotoxic shock.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..16
 FT /note= "hFas antigen signal peptide"
 FT Protein 30..376
 FT /note= "hFas (nd29) protein"
 FT
 XX WO9818487-A1.
 PN
 XX
 PD 07-MAY-1998.
 XX
 XX 31-OCT-1997; 97WO-JP03978.
 XX
 XX 26-SEP-1997; 97JP-0262521.
 PR
 XX 31-OCT-1996; 96JP-0290459.
 PR
 XX 27-DEC-1996; 96JP-0351718.
 PR
 XX (MOCH) MOCHIDA PHARM CO LTD.
 PA (OSAB-) OSAKA BIOSCIENCE INST.
 PA
 XX
 XX Nagata S, Suda T, Yatomi T;
 PI
 XX
 XX WPI; 1998-271925/24.
 DR
 XX N-PSDB; AAV34430.
 DR
 XX
 XX Use of Fas antagonist for treatment and prevention of
 PT apoptosis-related diseases - such as heart or kidney failure,
 PT graft-versus-host disease or liver disease
 PT
 XX Examples; Fig 5-9; 86pp; Japanese.
 PS
 XX
 XX This represents the antigenic peptide hFas (nd29) containing the Fc
 CC region. The invention provides the use of Fas antagonist as an agent for
 CC the treatment and prevention of apoptosis-related diseases. The Fas
 CC antagonist can be a partial Fas antigen peptide containing the
 CC extracellular part of the protein, but lacking the signal sequence, an
 CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is
 CC preferably a humanised antibody. The Fas antagonist is used in the
 CC treatment and prevention of diseases such as myocardial infarction, heart
 CC failure, ischemic heart disease, acute kidney failure, graft-versus-host
 CC disease, ischemic restenosis of the heart, liver or kidney, and
 CC endotoxic shock, and also as an organ preservative in transplantation.
 CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand
 CC system.
 CC
 XX
 XX Sequence 376 AA;

Query Match 74.3%; Score 1247; DB 19; Length 376;
 Best Local Similarity 65.4%; Pred. No. 6.6e-87;
 Matches 249; Conservative 10; Mismatches 32; Indels 90; Gaps 8;
 QY 14 VLGLWLAL---LRSVA-----GEQ-----APGTAPCSRGS 40
 DB 1 mlglwtllpvltsvatnleghlhdgqfchkcpcpgerkardctvngdepdcvpcqegk 60
 QY 41 SWSADL---DKCMDC-----ASCRARP-----HSDFLCLG 67
 DB 61 eytdkafhfskrrcrlcddeghgleveinctrtqntkrcrkpnffcnstvtcehdpcctkc 120

QY 68 -----AAAPPAPFRLWRSCDKTHTCCPAPAEAGAPSVFLPFPK 108
| :|||||
Db 121 ehgiikec|tsntckee|srnep-----kscdk|htccp|papellg|psvflfpk 175
QY 109 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVL 168
| :|||||
Db 176 pkdtlmi|rtpevtcvvdvsh|edpevkfnw|yvdgve|vhnaktk|preeqynst|rvvsvl 235
QY 169 TVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLT 228
| :|||||
Db 236 tvl|hqdw|lngkeyck|vsnkal|papi|ektisk|akgq|prep|qv|t|pp|sre|emtk|nqv|sl|t 295
QY 229 CLVKGFPYSDIAVESNGQPENNYKTTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCS 288
| :|||||
Db 296 clv|kgf|psd|ia|vesng|qpenn|ykt|tp|pv|ld|sd|gs|ff|ly|sk|lt|vd|k|sr|w|qq|gn|vf|sc|s 355
QY 289 VNHEALHNHYTQKSLSLSPGK 309
| :|||||
Db 356 vn|heal|hn|hyt|qks|ls|sl|spgk 376
RESULT 10
AAV15239
ID . AAV15239 standard; Protein; 450 AA.
XX
AC AAV15239;
XX
DT 04-NOV-1999 (first entry)
XX
DE ULBP1-Fc fusion polypeptide amino acid sequence.
XX
KW cell surface glycoprotein; lymphoma; assay; retinitis pigmentosa;
KW diabetes mellitus; progressive pseudo rheumatoid arthropathy;
KW muscular dystrophy; congenital merosin-deficiency; cardiomyopathy.
KW fusion polypeptide.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO9931241-A1.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-US27048.
XX
PR 15-JUL-1998; 98US-0092946.
PR 17-DEC-1997; 97US-0069857.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Cosman DJ, Fanslow WC, Mullberg JH;
XX
DR WPI; 1999-493777/41.
XX
PT DNA encoding human B cell lymphocyte cell surface glycoproteins
XX
PS Disclosure; Page 29; 104pp; English.
XX
This is the amino acid sequence of the ULBP1-Fc fusion polypeptide. The
ULBP-1 and ULBP-2 polypeptides can be expressed as Fc fusion proteins
using an Fc mutain to provide fused polypeptides, such as ULBP1-Fc.
The polypeptides ULBP-1 and -2 bind UL16-Fc. In addition, they bind to a
number of human cell types, including mitogen-stimulated human T cells
and natural killer (NK) cells. ULBP-Fc proteins bind to K299 cells, an
anaplastic lymphoma. The ULBP proteins can therefore be used as markers
to detect cancer, to enhance IFN- gamma production, NK cell proliferation
and CTL activity, to purify proteins and measure their activity. The
polypeptides and their fragments can also be used as delivery and
therapeutic agents, for rational drug design, as research reagents,
controls for peptide fragmentation, molecular weight/isoelectric focusing
markers, identification of unknown proteins and also for preparation of
antibodies.

The antibodies can be used in assays to detect the presence of ULBP
proteins, in vitro or in vivo, as well as for use in purification of ULBP
proteins. The ULBP DNA sequences can be used as probes to identify
homologues, to identify human chromosome number 6, to map and identify
genes, especially associated with certain diseases, syndromes or other
conditions on chromosome 6, as single-stranded sense or antisense
oligonucleotides, to inhibit expression of ULBP polypeptides, to help
detect defective genes in an individual and for gene therapy.
Diseases, syndromes and conditions associated with human chromosome
6 include retinitis pigmentosa (6q14-q21), Diabetes mellitus
(insulin-dependent diabetes mellitus) (6q21), progressive pseudorheumatoid arthropathy of
childhood (6q22), Muscular dystrophy (congenital merosin-deficient)
(6q22-q23), and cardiomyopathy (dilated) (6q23).
XX Sequence 450 AA;
SQ
Query Match 74.3%; Score 1247; DB 20; Length 450;
Best Local Similarity 96.3%; Pred. NO. 8.3e-87;
Matches 232; Conservative 2; Mismatches 5; Indels 2; Gaps 1;
QY 71 PP--APFRLWRSCDKTHTCCPAPAEAGAPSVFLPFPKPKDTLMISRTPEVTCVVVDV 128
| :|||||
Db 210 pps|lap|gtt|qpr|scd|k|ht|ccp|pap|eag|aps|v|fl|p|p|k|p|k|d|t|l|m|s|r|t|p|e|v|t|c|v|v|d|v 269
QY 129 SHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNK 188
| :|||||
Db 270 sh|ed|pe|vk|fn|w|y|d|g|v|e|v|h|n|a|k|t|k|p|r|e|e|q|y|n|s|t|y|r|v|v|s|v|l|t|v|l|h|q|d|w|l|n|g|e|y|c|k|v|s|n|k 329
QY 189 ALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQ 248
| :|||||
Db 330 alp|api|ekt|isk|akg|q|p|r|e|p|q|v|y|t|l|p|p|s|r|e|e|m|t|k|n|q|v|s|l|t|c|l|v|k|g|f|y|p|s|d|ia|v|e|w|e|s|n|g|q 389
QY 249 PENNYKTTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVNHEALHNHYTQKSLSLSPG 308
| :|||||
Db 390 pen|ny|k|t|t|p|p|v|l|d|s|d|g|s|f|f|l|y|s|k|l|t|v|d|k|s|r|w|q|q|n|v|f|s|c|s|v|n|h|e|a|l|h|n|y|t|q|k|s|l|s|l|s|p|g 449
QY 309 k 309
Db 450 k 450
RESULT 11
AAV15240
ID AAV15240 standard; Protein; 453 AA.
XX
AC AAV15240;
XX
DT 04-NOV-1999 (first entry)
XX
DE ULBP2-Fc fusion polypeptide amino acid sequence.
XX
KW cell surface glycoprotein; lymphoma; assay; retinitis pigmentosa;
KW diabetes mellitus; progressive pseudo rheumatoid arthropathy;
KW muscular dystrophy; congenital merosin-deficiency; cardiomyopathy.
KW fusion polypeptide.
XX
OS Homo sapiens.
OS
XX
PN WO9931241-A1.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-US27048.
XX
PR 15-JUL-1998; 98US-0092946.
PR 17-DEC-1997; 97US-0069857.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Cosman DJ, Fanslow WC, Mullberg JH;
XX
DR WPI; 1999-493777/41.

XX DNA encoding human B cell lymphocyte cell surface glycoproteins
PT Disclosure; Page 29; 104pp; English.
XX
XX This is the amino acid sequence of the ULBP2-Fc fusion polypeptide. The
XX ULBP-1 and ULBP-2 polypeptides can be expressed as Fc fusion proteins
XX using an Fc mutin to provide fused polypeptides, such as ULBP2-Fc.
XX The polypeptides ULBP-1 and -2 bind ULBP-1. In addition, they bind to a
XX number of human cell types, including mitogen-stimulated human T cells
XX and natural killer (NK) cells. ULBP-Fc proteins bind to K299 cells, an
XX anaplastic lymphoma. The ULBP proteins can therefore be used as markers
XX to detect cancer, to enhance IFN- gamma production, NK cell proliferation
XX and CTL activity, to purify proteins and measure their activity. The
XX polypeptides and their fragments can also be used as delivery and
XX therapeutic agents, for rational drug design, as research reagents,
XX controls for peptide fragmentation, molecular weight/isoelectric focusing
XX markers, identification of unknown proteins and also for preparation of
XX antibodies.
XX The antibodies can be used in assays to detect the presence of ULBP
XX proteins, in vitro or in vivo, as well as for use in purification of ULBP
XX proteins. The ULBP DNA sequences can be used as probes to identify
XX homologues, to identify human chromosome number 6, to map and identify
XX genes, especially associated with certain diseases, syndromes or other
XX conditions on chromosome 6, as single-stranded sense or antisense
XX oligonucleotides, to inhibit expression of ULBP polypeptides, to help
XX detect defective genes in an individual and for gene therapy.
XX Diseases, syndromes and conditions associated with human chromosome
XX 6 include Retinitis pigmentosa (6q14-q21), Diabetes mellitus
XX (insulin-dependent) (6q21), progressive pseudorheumatoid arthropathy of
XX childhood (6q22), Muscular dystrophy (congenital merosin-deficient)
XX (6q22-q23), and cardiomyopathy (dilated) (6q23).
XX
SQ Sequence 453 AA;

Query Match 74.3%; Score 1247; DB 20; Length 453;
Best Local Similarity 91.5%; Pred. No. 8.3e-87;
Matches 236; Conservative 3; Mismatches 9; Indels 10; Gaps 2;
QY 62 DFCLG-----CAAPPA---PRLLRSCDKTHCPCPAPEAGAPSVFLPPKPKD 111
Db 196 dflmgdstleapagaplanstgtqlrrscdkthtccpapeagapsvflppkpkd 255
QY 112 TLMSTRPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREQYNSTRYVSVLTVL 171
Db 256 tlmstrpevtcvvdvshedpevkfnwydgvvhnaktkpreeqynstryrsvltvl 315
QY 172 HQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREMTKNOVSLTCLV 231
Db 316 hqdwlngkeykckvsnkalpapietkiskakgprepqvyltppsdeltknqysltclv 375
QY 232 KGYFSPDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFNCSVMH 291
Db 376 kgfypsdiavwesngqpennnyktpvpldsdgsfflyskltvdksrwwqgnvfscvmh 435
QY 292 EALHNHYTQKSLSLSPCK 309
Db 436 ealhnhytqkslsispck 453
RESULT 12
AAW14764
ID AAW14764 standard; Protein; 424 AA.
XX
AC AAW14764;
DT 11-JUN-1997 (first entry)
XX Human soluble kit ligand-IgG fusion protein.
XX
XX Kit ligand; c-kit proto-oncogene; cytokine; growth factor;
KW haematopoietic cell; cell proliferation; stem cell; anaemia;
KW

KW thrombocytopaenia; therapy; IgG1.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
XX Peptide 1..25
FT /label= "Sig_peptide"
FT /note= "KL signal peptide"
FT Protein 26..424
FT /label= Mat_protein
FT /note= "human KL-IgG fusion"
PN W09526199-A1.
XX
XX 05-OCT-1995.
XX
XX 28-MAR-1995; 95WO-US03866.
XX
XX 28-MAR-1994; 94US-0220379.
XX (CYTO-) CYTOMED INC.
XX
XX Lobell RB, Nocka KH;
XX
XX WPI: 1995-351198/45.
XX N-PSDB; AAT63109.
XX
XX Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit
XX increased activity in promoting cell proliferation
XX
XX Claim 10; Page 43-44; 88pp; English.
XX
XX A fusion protein (AAW14764) between human soluble kit ligand (KL)
XX (see also AAW14761) and a human IgG1 heavy chain can be transiently
XX expressed in COS cells transfected with a human KL-Ig cDNA
XX construct (AAT63109) in vector CDM8; a corrected KL-Ig construct
XX (AAW14765) has also been prepd. KL-Ig can also be produced as a
XX dimer stabilised by intermolecular disulphide bonds or a peptide
XX linker. The stabilised KL-Ig dimers have a more favorable cell
XX proliferation/mast cell activation ratio than native KL and can
XX stimulate haematopoietic recovery or stem cell/progenitor cell
XX mobilisation with less toxicity.
XX
SQ Sequence 424 AA;

Query Match 74.1%; Score 1243; DB 16; Length 424;
Best Local Similarity 84.6%; Pred. No. 1.5e-86;
Matches 237; Conservative 6; Mismatches 25; Indels 12; Gaps 2;
QY 30 APGTAPCSRGSNSADLDKCMDCASCRRPHSDFCLGCAAPAPPAPFLLRSCDKHTCP 89
Db 157 asetsdcvssstlspekds-----rvsvtkpfpmlppvaadpep-----kscdkhtcp 204
QY 90 PCAPAEAGAPSVFLPPKPKDLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVHNA 149
Db 205 pcapaelggpsvflfppkpkdmlisrtpevtcvvdvshedpevkfnwydgvvhn 264
QY 150 KTRPREEQYNSTRYVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREP 209
Db 265 ktkpreeqynstryrsvsvltvlhqdwlngkeykckvsnkalpapietkiskakgqprep 324
QY 210 VYTLPPSREMTKNOVSLTCLVKGFPSPDIAVWESNGQPENNYKTPPVLDSDGSFFLY 269
Db 325 vytlpsrdeitknqysltclvkgfypsdiavwesngqpennnyktpvpldsdgsffly 384
QY 270 SKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSLSPCK 309
Db 385 skltvdksrwwqgnvfscvmhealhnhytqkslsispck 424
RESULT 13
AAW14765

CC They are also useful in screening assays for (ant)agonists of HER3
CC and HER4, and bind to cells that express HER4 e.g. MDA-MB-453 human
CC breast cancer cells, inducing expression of intercellular adhesion
CC molecule-1.
XX
SQ Sequence 375 AA;

Query Match 73.9%; Score 1239.5; DB 18; Length 375;
Best Local Similarity 93.9%; Pred. No. 2.5e-86;
Matches 229; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY 66 GCAAAPAPFRLWRSCDKTHTPCPCPAEAGAPSVFLPPPKDITLMSRTPEVTCVV 125
Db 137 gfsgdpep-----kscdkthtccpapefegapsvflfppkpkdtlmisrtpevtcv 191
QY 126 VDVSHEDEPKFNNYVDGVEVHNNAKTPREEQYNSTYRVSVTLVHQDLNKEKCKV 185
Db 192 vdvshdedevkfnwydgvvevhnaktpreeqynstyrvvsvtlvhdwlngkeykcv 251
QY 186 SNKALPAPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEWES 245
Db 252 snkalpapielktiskagqpreqvyltppsrdeitknqvsitclvkgyfypsdiawes 311
QY 246 NGOPENNYKTPPVLDSDGFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSL 305
Db 312 ngqpennyktpvpldsgdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkls 371
QY 306 SPQK 309
Db 372 spgk 375

RESULT 15
AAV97175
ID RAY97175 standard; Protein; 488 AA.
XX
AC AAV97175;
XX
DT 04-DEC-2000 (first entry)
XX
DE Human FGF-RI Extracellular domain-Ig Fc fusion protein 6.
XX
KW FGF-R; fibroblast growth factor receptor; extracellular domain; IgG1;
KW immunoglobulin; G1; oligomerization domain; Fc region; fusion protein;
KW inhibitor; dimer; antagonist; cytostatic; anti-diabetic; vulnerary;
KW ophthalmological; anti-proliferative.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Peptide 1..21
FT Domain /label= FGF-RI_signal_peptide
FT Domain 22..257
FT Domain /label= FGF-RI_extracellular_domain
FT Domain /note= "the Ig I segment and acid box are deleted"
FT Domain 59..111
FT Domain /label= Ig_II_segment
FT Domain 157..222
FT Region /label= Ig_III_segment
FT Region 258..488
FT Region /label= Human_IgG1_Fc_region
FT Region /note= "Contains hinge region and domains CH2 and CH3"
FT Misc-difference 276
FT /label= L276E
FT /note= "this mutation decreases the affinity of the Fc
FT portion for Fc receptors"
FT
FT Misc-difference 376
FT /label= p376S
FT /note= "this mutation decreases the affinity of the Fc
FT portion for complement"
FT
FT
XX WC2000046380-A2.

XX 10-AUG-2000.
XX
PF 07-FEB-2000; 2000WO-US03166.
XX
PR 08-FEB-1999; 99US-0119002.
XX
PA (CHIR) CHIRON CORP.
XX
PI Kavanaugh WM, Ballinger M;
XX
DR WPI; 2000-514961/46.
XX
PT N-PSDB; AAA52132.
XX
PT New polypeptide comprising a fibroblast growth factor receptor
PT extracellular domain fused to a heterologous oligomerization domain for
PT treating FGF-, angiogenesis-, or FGF receptor-mediated disorders
XX
XX Claim 14; Page 68-69; 70pp; English.
XX
CC Novel fusion protein constructs comprise a fibroblast growth factor (FGF)
CC receptor (FGF-R) extracellular domain (ECD) lacking the immunoglobulin
CC (Ig) I segment fused to a heterologous oligomerization domain that
CC comprises an immunoglobulin Fc region, hinge region, CH1, CH2, CH3 or CH4
CC region, or light chain of an immunoglobulin molecule, or a peptide with a
CC leucine zipper motif. The Ig I segment is not necessary for binding of
CC acidic FGF and basic FGF (bFGF). The Ig I deletion further increases the
CC affinity for aFGF and heparin, protects the core of the molecule from
CC proteolysis, and abrogates the heparin requirement for aFGF binding. The
CC new fusion polypeptides are better FGF inhibitors than FGF-R monomer
CC proteins. The FGF-R-Ig Fc fusion dimers are active as FGF antagonists at
CC subnanomolar concentrations and were 20-fold more potent than the FGF-R
CC monomer protein as competitors of bFGF binding to immobilized FGF-Rs. The
CC fusion constructs are useful to treat FGF-, angiogenesis- or
CC FGF-R-mediated disorders, such as tumorigenesis (e.g. bladder, breast,
CC lung, rectal, testis and cervical tumours), neovascularization (e.g.
CC diabetic retinopathy, neovascular glaucoma, wound healing and corneal
CC scarring) and hyper-proliferation of vascular smooth muscle cells (e.g.
CC postangioplasty and postatherectomy restenosis).
XX
SQ Sequence 488 AA;

Query Match 73.8%; Score 1239; DB 21; Length 488;
Best Local Similarity 81.1%; Pred. No. 3.7e-86;
Matches 240; Conservative 9; Mismatches 23; Indels 24; Gaps 4;

QY 22 LRSVAGEQAPGTAPCSRSG-----SSWSADLDKCMDCASCRARPHSDFLCGCAAAPPA 73
Db 209 lrnvstfeda-geytcclagnsigishhsawitvie-----aleerp-----avmts 252
QY 74 PFRLLWRSCDKTHTPCPCPAEAGAPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDP 133
Db 253 plylepkscdkthtccpapelegpsvflfppkpkdtlmisrtpevtcvvvdvshep 312
QY 134 EVFNNVVDGVEVHNNAKTPREEQYNSTYRVSVTLVHQDLNKEKCKVSNKALPAP 193
Db 313 evkfnwydgvvevhnaktpreeqynstyrvvsvtlvhdwlngkeykcvsnkalpas 372
QY 194 IEKTISKAGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNY 253
Db 373 iektiskagqpreqvyltppsrdeitknqvsitclvkgyfypsdiawesngqpenny 432
QY 254 KTPPVLDSDGFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPGK 309
Db 433 ktppvldsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqklsispgk 488

Search completed: March 11, 2002, 15:14:13
Job time: 97 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:53 ; Search time 31.92 Seconds
(without alignments)
672.971 Million cell updates/sec

Title: US-09-742-454A-7_COPY_28_309

Perfect score: 1551

Sequence: 1 EQAPGTAPCGSGSSWSADLD.....MHEALHNHYTKSLSLSPGK 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1230	79.3	374	2	S69339
2	1224	78.9	255	4	S31866
3	1224	78.9	330	1	GHU
4	1171.5	75.5	377	2	A23511
5	1171	75.5	234	2	PT0207
6	1169.5	75.4	377	2	A60764
7	1158.5	74.7	326	1	G2HU
8	1146.5	73.9	289	1	G3HUI
9	1138	73.4	327	1	G4HU
10	927	59.8	328	2	I47159
11	923	59.5	328	2	I47160
12	910.5	58.7	323	1	GHRB
13	907	58.5	277	2	I47162
14	894	57.6	328	2	I47158
15	892	57.5	328	2	I47161
16	876	56.5	329	1	G2GP
17	859.5	55.4	470	2	S22080
18	842	54.3	472	2	S31459
19	840	54.2	308	2	C30554
20	840	54.2	444	2	PC4436
21	838	54.0	326	2	P50017
22	836	53.9	329	1	G3MSC
23	833	53.7	324	1	G1MS
24	828	53.4	393	1	G1MSM
25	825	53.2	333	2	PS0018
26	825	53.2	398	1	G3MSM
27	817	52.7	329	2	S00847
28	804.5	51.9	322	2	PS0019
29	799.5	51.5	330	1	G2MSA

30	799.5	51.5	469	2	S37483	Ig gamma-2a chain
31	794.5	51.2	399	1	G2MSAM	Ig gamma-2a chain
32	794	51.2	335	1	G2MSAB	Ig gamma-2a chain
33	786.5	50.7	474	1	G2MS11	Ig gamma-2b chain
34	786	50.7	446	2	S40295	Ig gamma-2a chain
35	781.5	50.4	405	1	G2MSBM	Ig gamma-2a chain
36	776.5	50.1	327	2	S06611	Ig gamma-2 chain C
37	762.5	49.2	475	2	S01321	Ig gamma-2b chain
38	708	45.6	180	2	I46732	Ig gamma heavy chain
39	583.5	37.6	249	2	S69340	Ig heavy chain VH1
40	577	37.2	152	2	S4236	Ig gamma-1 chain C
41	568.5	36.7	218	2	A36040	Ig heavy chain V-1
42	410	26.4	572	2	B46529	Ig y heavy chain (
43	367	23.7	453	2	S37768	Ig mu chain C regi
44	366.5	23.6	388	1	EHMS	Ig epsilon chain C
45	365	23.5	391	1	MHHUBT	Ig mu heavy chain

ALIGNMENTS

RESULT 1

S69339

Ig heavy chain V region precursor - human

C:Species: Homo sapiens (man)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000

C:Accession: S69339; S72664

R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.

A:Reference number: S69339; MUID:95262687

A:Accession: S69339

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <KHA>

A:Cross-references: EMBL:X81695

R:Khamilichi, A.A.

submitted to the EMBL Data Library, September 1994

A:Reference number: S72664

A:Accession: S72664

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140,'C',142-374 <KH2>

A:Cross-references: EMBL:X81695

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 79.3%; Score 1230; DB 2; Length 374;
Best Local Similarity 98.3%; Pred. No. 8.1e-82;
Matches 226; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDTHTCPPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112

Db 145 KSCDTHTCPPCPAPEAGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 204

QY 113 YVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSKALPAPIEKTIS 172

Db 205 YVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSKALPAPIEKTIS 264

QY 173 KAKGPREFQVYTLPPSREEMTKNQVSLTCLVKGYFSPDAVWEESNGQPPENNYKTPPV 232

Db 265 KAKGPREFQVYTLPPSREEMTKNQVSLTCLVKGYFSPDAVWEESNGQPPENNYKTPPV 324

QY 233 LDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSPGK 282

Db 325 LDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSPGK 374

RESULT 2

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C;Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C;Accession: S31866

R;Filpula, D.

submitted to the EMBL Data Library, February 1993

A;Description: Screening method for protein-protein interactions of cloned gene products.

A;Reference number: S31866

A;Accession: S31866

A;Molecule type: mRNA

A;Residues: 1-255 <FIL>

A;Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C;Keywords: immunoglobulin

F;1-22/Region: Escherichia coli outer membrane protein A precursor

F;23-255/Region: human Ig gamma-1 chain C region

Query Match 78.9%; Score 1224; DB 4; Length 255;
Best Local Similarity 97.4%; Pred. No. 1.4e-81;
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPCPAPEAEGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112

Db :|||||

Db 26 KSCDKTHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 85

QY 113 YVDGVEVHNKTKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIS 172

Db :|||||

Db 86 YVDGVEVHNKTKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIS 145

QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 232

Db :|||||

Db 146 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 205

QY 233 LDSGSEFLYSLKLTVDKSRQOQGVNFVCSVMHEALHNHYTQKSLSLSPGK 282

Db :|||||

Db 206 LDSGSEFLYSLKLTVDKSRQOQGVNFVCSVMHEALHNHYTQKSLSLSPGK 255

RESULT 3

IGHU

Ig gamma-1 chain C region - human

C;Species: Homo sapiens (man)

C;Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999

C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R;Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A;Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.

A;Reference number: A93433; MUID:82274238

A;Accession: A93433

A;Molecule type: DNA

A;Residues: 1-330 <ELL>

A;Cross-references: EMBL:Z17370

A;Note: this sequence has the Glm(17) allotypic marker, 97-Lys, and the Glm(1) markers,

A;Note: Lys-330 is removed after translation

R;Harris, L.J.

submitted to the EMBL Data Library, October 1992

A;Reference number: S33904

A;Accession: S36861

A;Molecule type: DNA

A;Residues: 2-330 <HAR>

A;Cross-references: EMBL:Z17370

R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of a

A;Reference number: S33887; MUID:83001943

A;Accession: S33887

A;Molecule type: DNA

A;Residues: 88-113;235-330 <TAK>

A;Cross-references: EMBL:Z17370

R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,

Biochemistry 9, 3161-3170, 1970

A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen

A;Reference number: A90563; MUID:71064024

A;Contents: myeloma protein Eu

A;Accession: B90563

A;Molecule type: protein

A;Residues: 1-96,'R',98-135 <CUN>

A;Note: this sequence has the Glm(3) marker, 97-Arg

R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se

A;Reference number: A90564; MUID:71064025

A;Contents: Eu

A;Accession: A90564

A;Molecule type: protein

A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2

A;Note: this sequence has the Glm(non-1) markers, 239-Glu and 241-Met

R;Ponstingl, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni

igen Primaerstruktur.

A;Reference number: A91668; MUID:77070269

A;Contents: myeloma protein Nie

A;Accession: B91668

A;Molecule type: protein

A;Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E'

A;Note: this sequence has the Glm(17) and Glm(1) markers

R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983

A;Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1

A;Reference number: A91723; MUID:83289131

A;Contents: myeloma protein KOL; disulfide bonds

A;Accession: A91723

A;Molecule type: protein

A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <

A;Note: this sequence has the Glm(3) and Glm(non-1) markers

R;Gall, W.E.; Edelman, G.M.

Biochemistry 9, 3188-3196, 1970

A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disul

A;Reference number: A90565; MUID:71064027

A;Contents: annotation; disulfide bonds

R;Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976

A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunog

enbromide cleavage products, and the disulfide bridges.

A;Reference number: A91667; MUID:77070267

A;Contents: annotation; disulfide bonds

C;Genetics:

A;Gene: GDB:IGHG1

A;Cross-references: GDB:120085; OMIM:147100

A;Map position: 14q32.33-14q32.33

A;Introns: 99/1; 114/1; 224/1

C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C;Superfamily: immunoglobulin C region; immunoglobulin homology

C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F;20-85/Domain: immunoglobulin homology <IMI>

F;137-206/Domain: immunoglobulin homology <IM2>

F;243-310/Domain: immunoglobulin homology <IM3>

F;27-83,144-204,250-308/Disulfide bonds: #status experimental

F;103/Disulfide bonds: interchain (to light chain) #status experimental

F;109,112/Disulfide bonds: interchain (to heavy chain) #status experimental

F;180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 78.9%; Score 1224; DB 1; Length 330;
Best Local Similarity 97.4%; Pred. No. 1.9e-81;
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPCPAPEAEGAPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 112

Db :|||||

Db 101 KSCDKTHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 160

QY 113 YVDGVEVHNKTKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIS 172

Db :|||||

Db 161 YVDGVEVHNKTKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIS 220

QY 173 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPV 232

Ig gamma-2 chain C region - human
G2HU
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1981 #sequence_rev
C;Accession: A93906; A92809; A907

R;Ellison, J.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
A;Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con
A;Reference number: A93506; MUID:82157621
A;Accession: A93506
A;Molecule type: DNA
A;Residues: 1-326 <ELL>
A;Cross-references: GB:V00554; GB:J00230; NID:g32759; PIDN:CAB58438.1; PID:g6066056
A;Note: Lys-326 is probably removed posttranslationally
R;Wang, A.C.; Tung, E.; Fudenberg, H.H.
J. Immunol. 125, 1048-1054, 1980
A;Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f
A;Reference number: A92809; MUID:81007873
A;Contents: myeloma protein Til
A;Accession: A92809
A;Molecule type: protein
A;Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>
A;Note: Trp-156 is at or near the complement-binding site
R;Connell, G.E.; Parr, D.M.; Hofmann, T.
Can. J. Biochem. 57, 758-767, 1979
A;Title: The amino acid sequences of the three heavy chain constant region domains of a
A;Reference number: A90752; MUID:80001357
A;Contents: myeloma protein Zie
A;Accession: A90752
A;Molecule type: protein
A;Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-
A;Note: this sequence has since been revised
R;Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A;Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
A;Reference number: A93132; MUID:80114419
A;Contents: Zie
A;Accession: A93132
A;Molecule type: protein
A;Residues: 238-275 <HOP>
R;Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A;Reference number: A94591
A;Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268
A;Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
ned
R;Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A;Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.
A;Reference number: A90253; MUID:72033500
A;Contents: annotation; myeloma protein Sa, disulfide bonds
R;Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A;Title: Structural studies of immunoglobulin G.
A;Reference number: A93157; MUID:69064124
A;Contents: annotation; Sa, disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG2
A;Cross-references: GDB:I19338; OMIM:147110
A;Map position: 14q32.33-14q32.33
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kd
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;20-85/Domain: immunoglobulin homology <IMI>
F;133-202/Domain: immunoglobulin homology <IM2>
F;239-306/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;27-83,140-200,246-304/Disulfide bonds: #status experimental
F;102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.7%; Score 1158.5; DB 1; Length 326;
Best Local Similarity 68.0%; Pred. No. le-76;
Matches 227; Conservative 8; Mismatches 22; Indels 77; Gaps 5;
7 APCSRGSSWSADLKCDCASCARPHSDFCLGCAAP--PAPPRLLWRS----- 54

||||| :| | ||| :| | :| | :| |
Db 12 APCSRSTS-----ESTALGLCKVDYFPEPVTVSNWNSGALTSVHTF 53
Qy 55 -----CDKTH-----TCPPCPAPE 68
Db 54 PAVLQSSGLYSLSSVTVVPSNSFTQYTVTNDVHKPSNTKVDKTVKRCCKVCEPCPPAPP 113
Qy 69 AEGAPSVFLPPPKDKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDGVEVHNNAKTKPRE 128
Db 114 VAG-PSVFLPPPKDKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDGVEVHNNAKTKPRE 172
Qy 129 EQYNSTYRVVSVLTVLHQDWLNGKEYCKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPP 188
Db 173 EQNSTERVVSVLTWVHQDWLNGKEYCKCKVSNKGLPAPIEKTIISKAKGQPREPQVYTLPP 232
Qy 189 SREEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFSLSKLTVD 248
Db 233 SREEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFSLSKLTVD 292
Qy 249 KSRWQOQGVFSCSVNHEALHNHYTQKSLSLSPGK 282
Db 293 KSRWQOQGVFSCSVNHEALHNHYTQKSLSLSPGK 326
RESULT 8
G3HUW1
Ig gamma-3 heavy chain disease proteins - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C;Accession: A90442; A92219; A90198; A93915; A02149
R;Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A;Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 hea
A;Reference number: A90442; MUID:81021548
A;Contents: heavy chain disease protein Wis
A;Accession: A90442
A;Molecule type: protein
A;Residues: 1-289 <PRA>
A;Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interch
A;Note: this protein lacks most of the V region and all of the CH1 region. Residue 12
R;Michaelson, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A;Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplicat
A;Reference number: A92219; MUID:77118561
A;Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protei
A;Accession: A92219
A;Molecule type: protein
A;Residues: 12-97 <MIC>
A;Note: the hinge region in gamma-3 chains is about four times as long as in other ga
idue segment (12-28)
A;Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form in
R;Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A;Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of t
A;Reference number: A90198; MUID:77021516
A;Contents: heavy chain disease protein Zuc, partial sequence corresponding to residu
A;Accession: A90198
A;Molecule type: protein
A;Residues: 59-125, 'BB', 128-226, 228-289 <WOL>
A;Note: this protein lacks most of the V region, all of the CH1 region, and part of t
R;Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood,
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A;Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deleti
A;Reference number: A93915; MUID:82247835
A;Contents: heavy chain disease protein Omm
A;Accession: A93915
A;Molecule type: mRNA
A;Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-
A;Note: a carboxyl-terminal Lys is removed posttranslationally
A;Note: this sequence may represent an allelic form or another gamma chain subclass
C;Comment: The heavy chain disease protein Wis is shown.
C;Genetics:

A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglytamic acid
F:203-270/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 73.9%; Score 1146.5; DB 1; Length 289;
Best Local Similarity 81.1%; Pred. No. 6.6e-76;
Matches 214; Conservative 14; Mismatches 33; Indels 3; Gaps 1;

QY 21 KCMDCASCARPHSDFC---LCCAAAPAPFRLLRSCDTHTCPCPAPEAGAPSVFL 77
DB 26 RCPEPKSCDTPPCPCPEPKSCDTPPCPEPKSCDTPPCPCPAPEAGAPSVFL 85
QY 78 FPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDGVEVHNKTRPREQYNSTRYV 137
DB 86 FPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDGVEVHNKTRPREQYNSTRYV 145
QY 138 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQ 197
DB 146 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQ 205
QY 198 VSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNV 257
DB 206 VSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNV 265
QY 258 FSCSVWHEALHNHYTQKSLSLSPG 281
DB 266 FSCSVWHEALHNHYTQKSLSLSPG 289

RESULT 9
G4HU
Ig gamma-4 chain C region - human
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
C:Accession: A90933; A90249; A02150
R:Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A:Reference number: A90933; MUID:83157104
A:Accession: A90933
A:Molecule type: DNA
A:Residues: 1-327 <ELL>
A:Note: the sequence was determined from the germline gene
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant
A:Reference number: A90249; MUID:70207560
A:Accession: A90249
A:Molecule type: protein
A:Residues: 1-30; 81-326 <PIN>
C:Genetics:
A:Gene: GDB:IGHG4
A:Cross-references: GDB:119340; OMIM:147130
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 111/1; 221/1
A:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:99-110/Region: hinge
F:134-203/Domain: immunoglobulin homology <IM2>
F:240-307/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,141-201,247-305/Disulfide bonds: #status predicted
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.4%; Score 1138; DB 1; Length 327;
Best Local Similarity 67.1%; Pred. No. 3.1e-75;
Matches 224; Conservative 9; Mismatches 25; Indels 76; Gaps 4;

QY 7 APCSRGSSWSADLCKMDCASCARPHSDFCGLGCAAP--PAPFRLLR 54
DB 12 APCSRSTS-----ESTAAALGCLVKDYFPEPTVSWNSGALTSGVHTF 53
QY 55 -----CDKTH-----TCPPCPAPE 68
DB 54 PAVLOSSGLYSLSSVVTVFSSSLGTYTCNVHDKPSNTKVDKRVESKYGPCPCPAPE 113
QY 69 AGAPSVFLFPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDGVEVHNKTRPRE 128
DB 114 FLGGPSVFLFPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDGVEVHNKTRPRE 173
QY 129 EYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPP 188
DB 174 EGFNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPP 233
QY 189 SREEMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVD 248
DB 234 SOEEMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVD 293
QY 249 KSRWQQGNVFCVSHVHEALHNHYTQKSLSLSPG 282
DB 294 KSRWQQGNVFCVSHVHEALHNHYTQKSLSLSPG 327

RESULT 10
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845
A:Accession: I47159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:9433123; PID:AA52217.1; PID:9433124
C:Genetics:
A:Gene: IgG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 59.8%; Score 927; DB 2; Length 328;
Best Local Similarity 52.7%; Pred. No. 5.8e-60;
Matches 183; Conservative 35; Mismatches 43; Indels 86; Gaps 7;

QY 3 APGT-----APCSRGSWSADLCKMDCASCARPHSDFCGLGCAAP--PAPFRLLR 53
DB 1 APKTAPSVYPLAPCSRSTS-----GPNVALGCLASSYFPEPTVTWN 42
QY 54 S-----CDKTH----- 59
DB 43 SGALSSGVHTFSPVLPQSLGSLSSVMVTPASSLSKSYTCNVNHPATTKVDKRVGTKT 102
QY 60 --TCPPCPAPEAGAPSVFLFPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDG 117
DB 103 KPPCPICPACESPG-PSVFIEFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNQYVDG 161
QY 118 EYHNKTRPREQYNSTRYVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKG 177
DB 162 EYHTAQTRPEKQFNSTRYVSVLTPIHQDWLNKKEYCKVSNKALPAPIEKTISKAKG 221
QY 178 PREPQVYTLPPSREEMTKNQVSLTCLVKGFPYSDIAVEWESNGQ--PENNYKTTTPVLD 235

C:Accession: I47162
R:Kaskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A:Reference number: I47158; MUID:95015845
A:Accession: I47162
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <KAC>
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130
C:Genetics:
A:Gene: IgG4
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 58.5%; Score 907; DB 2; Length 277;
Best Local Similarity 72.1%; Pred. No. 1.3e-58;
Matches 165; Conservative 29; Mismatches 31; Indels 4; Gaps 3;

QY 58 THTCCPCP-APEAG-APSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 115
Db 49 TTKTKPCPCACEGPGSAFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 108
QY 116 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 175
Db 109 GVEVHTAQRKEEQFNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 168
QY 176 GQRPQVYTLPPSREEMTKNQVSLTCLVKGYFSDIAVEWESNGQ--PENNYKTTTPPV 233
Db 169 GQRPQVYTLPPSREEMTKNQVSLTCLVKGYFSDIAVEWESNGQ--PENNYKTTTPPV 233
QY 234 DSDGSEFLYSLKTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPGK 282
Db 229 DSDGSEFLYSLKTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPGK 277

RESULT 14
I47158
Ig gamma 1 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47158
R:Kaskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A:Reference number: I47158; MUID:95015845
A:Accession: I47158
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03778; NID:9433121; PIDN:AAA52216.1; PID:9433122
C:Genetics:
A:Gene: IgG1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 57.6%; Score 894; DB 2; Length 328;
Best Local Similarity 52.2%; Pred. No. 1.4e-57;
Matches 181; Conservative 30; Mismatches 50; Indels 86; Gaps 7;

QY 3 APGT-----APCSRGSSWSADLCKMDCASCARPHSDFCGLCAAAP--PAPFLLMR 53
Db 1 APKTAPSVYPLAPCGRDVS-----GPNVALGCLASSYFPEPVTMTWN 42
QY 54 S-----CDKTH----- 59
Db 43 SGALTSGVHTFSPVLPQSLGSLSSMVTVPASSLSKSYTCNVNHPATTKTKVDKRVGIHQ 102

QY 60 --TCPPCPAPEAGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGV 117
Db 103 PQTCTPCPCVAG-PSVFIFFPKPKDTLMISQTEPVTICVVVDVSKAEHAEVQFSWYVDGV 161

QY 118 EVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAKG 177
Db 162 EVHTAETRPKEQFNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAKG 221
QY 178 PREPQVYTLPPSREEMTKNQVSLTCLVKGYFSDIAVEWESNGQ--PENNYKTTTPPV 235
Db 222 SREPQVYTLPPSREEMTKNQVSLTCLVKGYFSDIAVEWESNGQ--PENNYKTTTPPV 281
QY 236 DSGSEFLYSLKTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPGK 282
Db 282 DSGSEFLYSLKTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPGK 328
RESULT 15
I47161
Ig gamma 3 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47161
R:Kaskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845
A:Accession: I47161
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03781; NID:9433127; PIDN:AAA52219.1; PID:9433128
C:Genetics:
A:Gene: IgG3
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 57.5%; Score 892; DB 2; Length 328;
Best Local Similarity 52.3%; Pred. No. 2e-57;
Matches 181; Conservative 32; Mismatches 49; Indels 84; Gaps 8;

QY 3 APGT-----APCSRGSSWSADLCKMDCASCARPHSDFCGLCAAAP--PAPFLLMR 53
Db 1 APKTAPSVYPLAPCGRDTS-----GPNVALGCLASSYFPEPVTMTWN 42
QY 54 S-----CDKTH----- 61
Db 43 SGALTSGVHTFSPVLPQSLGSLSSMVTVPASSLSKSYTCNVNHPATTKTKVDKRVGT 102
QY 62 -PPCP-APEAGA-PSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 118
Db 103 KPCCPCPCVAG-PSVFIFFPKPKDTLMISQTEPVTICVVVDVSKAEHAEVQFSWYVDGVE 162
QY 119 VINAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAKG 178
Db 163 VHTAETRPKEQFNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAKG 222
QY 179 REPQVYTLPPSREEMTKNQVSLTCLVKGYFSDIAVEWESNGQ--PENNYKTTTPPV 236
Db 223 REPQVYTLPPSREEMTKNQVSLTCLVKGYFSDIAVEWESNGQ--PENNYKTTTPPV 282
QY 237 GSEFLYSLKTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPGK 282
Db 283 GSEFLYSLKTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPGK 328

Search completed: March 11, 2002, 15:14:53
Job time: 137 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:52 ; Search time 19.78 Seconds
(without alignments)
522.724 Million cell updates/sec

Title: US-09-742-454A-7_COPY_28_309

Perfect score: 1551

Sequence: 1 EQAPGAPCSRGSWSADLD.....MHEALHNHYTKSLSPGK 282

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1224	78.9	330	1	GC1_HUMAN
2	1158.5	74.7	326	1	GC2_HUMAN
3	1151.5	74.2	290	1	GC3_HUMAN
4	1138	73.4	327	1	GC4_HUMAN
5	910.5	58.7	323	1	GC_RABIT
6	876	56.5	329	1	GC2_CAVPO
7	838	54.0	326	1	GC1_RAT
8	836	53.9	329	1	GC3_MOUSE
9	833	53.7	324	1	GC1_MOUSE
10	828	53.4	333	1	GC1M_MOUSE
11	825	53.2	333	1	GC2_RAT
12	825	53.2	398	1	GC3M_MOUSE
13	817	52.7	329	1	GC2_RAT
14	804.5	51.9	322	1	GCA_RAT
15	799.5	51.5	330	1	GCA_MOUSE
16	794.5	51.2	399	1	GCA_MOUSE
17	794	51.2	335	1	GCAB_MOUSE
18	786.5	50.7	336	1	GCBB_MOUSE
19	781.5	50.4	405	1	GCBM_MOUSE
20	367	23.7	454	1	MUC_HUMAN
21	366.5	23.6	421	1	EPC_MOUSE
22	365	23.5	391	1	MUCB_HUMAN
23	360.5	23.2	429	1	EPC_RAT
24	353.5	22.8	428	1	EPC_HUMAN
25	349	22.5	455	1	MUC_MOUSE
26	344	22.2	458	1	MUC_RABIT
27	339	21.9	476	1	MUCM_MOUSE
28	336	21.7	454	1	MUC_MESAU
29	334	21.5	479	1	MUCM_RABIT
30	333	21.5	457	1	MUC_SUNMU
31	331.5	21.4	450	1	MUC_CANFA
32	302.5	19.5	299	1	ALC_RABIT
33	290	18.7	438	1	HVC2_HETFR

34	288.5	18.6	446	1	MUC_CHICK
35	284.5	18.3	438	1	HVCS_HETFR
36	282	18.2	353	1	ALC1_HUMAN
37	279	18.0	353	1	ALC1_GORGO
38	275.5	17.8	461	1	HVCM_HETFR
39	274.5	17.7	340	1	ALC2_HUMAN
40	274	17.7	370	1	HVC1_HETFR
41	273.5	17.6	393	1	HVC3_HETFR
42	255	16.4	344	1	ALC_MOUSE
43	252.5	16.3	481	1	MUCM_ICTFU
44	178.5	11.5	105	1	LAC1_MOUSE
45	173	11.2	104	1	LAC2_RAT

ALIGNMENTS

RESULT 1
GC1_HUMAN STANDARD; PRT; 330 AA.
AC P01857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
GN IGHG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
Waxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RL Biochemistry 9:3161-3170(1970).
RN [3]
RP SEQUENCE OF 136-329 (EU).
RX MEDLINE=71064025; PubMed=5530842;
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RL Biochemistry 9:3171-3181(1970).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN NIE).
RX MEDLINE=77070269; PubMed=826475;
RA Pongstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [5]
RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71064027; PubMed=4923144;
RA Gall W.E., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. X:

RT Intrachain disulfide bonds.";
RL Biochemistry 9:3188-3196(1970).
RN [7]
RP DISULFIDE BONDS.
RX MEDLINE=77070267; PubMed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
"Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=81208100; PubMed=7236608;
RA Deisenhofer J.;
RT "Crystallographic refinement and atomic models of a human Fc fragment
RT and its complex with fragment B of protein A from Staphylococcus
RT aureus at 2.9- and 2.8-A resolution.";
RL Biochemistry 20:2361-2370(1981).
CC -!- MISCELLANEOUS: NIE HAS THE GLM(17) ALLOTYPIC MARKER, 97-K, & THE
CC GLM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GLM(3)
CC MARKER & THE GLM (NON-1) MARKERS.
CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
CC 35,116,198,269 & 272.
CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
CC 268-272.
CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
CC RESIDUES 198,267&272.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00228; AAC82527.1; ALT_INIT.
DR PIR: A02146; GHU.
DR PDB: 1FC1; 15-JUL-92.
DR PDB: 1FC2; 15-JUL-92.
DR MIM: 147100; -.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igc1; 2.
DR SMART: SM00410; Ig_like; 1.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 223 CH2.
FT DOMAIN 224 330 CH3.
FT DISULFID 27 83
FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT CARBOHYD 180 180
FT MOD_RES 330 330
FT VARIANT 97 97
FT VARIANT 239 239
FT VARIANT 241 241
FT STRAND 123 126
FT HELIX 130 134

FT TURN 136 137
FT STRAND 141 148
FT STRAND 158 162
FT TURN 163 164
FT STRAND 165 166
FT STRAND 175 178
FT STRAND 183 190
FT HELIX 193 197
FT TURN 198 199
FT STRAND 202 206
FT STRAND 215 219
FT STRAND 227 227
FT STRAND 230 234
FT HELIX 238 240
FT TURN 241 242
FT STRAND 245 256
FT STRAND 260 266
FT TURN 267 268
FT STRAND 269 270
FT STRAND 274 276
FT STRAND 280 281
FT TURN 283 284
FT STRAND 287 296
FT HELIX 297 301
FT TURN 302 303
FT STRAND 306 312
FT TURN 313 314
FT TURN 316 317
FT STRAND 320 324
SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;

Query Match 78.9%; Score 1224; DB 1; Length 330;
Best Local Similarity 97.4%; Pred. No. 1.le-86;
Matches 224; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 53 RSCDKTHTCPCPAPEAEGAPSVLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVAFNW 112
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 101 KSCDKTHTCPCPAPELGGPSVLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVAFNW 160

QY 113 YVDGVEVHNAKTPREQYNSTYRVYSLVTLHQDWLNGKEYCKVSNKALPAPIEKTIS 172
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 161 YVDGVEVHNAKTPREQYNSTYRVYSLVTLHQDWLNGKEYCKVSNKALPAPIEKTIS 220

QY 173 KAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 232
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 221 KAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 280

QY 233 LQSDGSGFFLYSKLTVDKSRWQQGNVSCFVSYMSHEALHNHYTQKSLSLSPGK 282
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 281 LQSDGSGFFLYSKLTVDKSRWQQGNVSCFVSYMSHEALHNHYTQKSLSLSPGK 330

RESULT 2
GC2_HUMAN
ID GC2_HUMAN STANDARD; PRT; 326 AA.
AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE IG GAMMA-2 CHAIN C REGION.
GN IGHG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
RT heavy chain constant region genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).


```
DR PIR: A02150; G4HU.
DR MIM: 147130; -.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; Igc1; 2.
DR SMART: SM00410; Igc1; 1.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 73.4%; Score 1138; DB 1; Length 327;
Best Local Similarity 67.1%; Pred. No. 4e-80;
Matches 224; Conservative 9; Mismatches 25; Indels 76; Gaps 4;

QY 7 APCSRGWSADLCKMDCASCRRPHSDFCIGCAAP--PAPFLLWRS----- 54
|||||:| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 12 APCSRSTS-----ESTAALGCLVKDYFPEPVTVSVNSGALTSQVHTF 53
QY 55 -----CDKTH-----TCPCPAPE 68
DB 54 PAVLOSSGLYSLSSVTVFPSSSLGKTKYTCNVDRHPSNTKVDKRVESYGGPCPCPAPE 113
QY 69 AEGAPSVLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPKRE 128
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 114 FLGGFVSLFPKPKDTLMISRTPEVTCVVDVSDQEDPEVQFNWYVDGVEVHNATKPKRE 173
QY 129 EGYNSTYRVSVTLVHODWLNKGYKCKVSNKALPAPIETKISAKAGPREPQVYTLPP 188
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 174 EGFNSTYRVSVTLVHODWLNKGYKCKVSNKGLPSSIEKTIKAKAGPREPQVYTLPP 233
QY 189 SREEMTKNOVSLTCLVKGYFSPSDIAVENESGQPNPNYKTTTPVLDSGGSFFLYSKLTVD 248
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 234 SGEEMTKNOVSLTCLVKGYFSPSDIAVENESGQPNPNYKTTTPVLDSGGSFFLYSKLTVD 293
QY 249 KSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 282
|||||:| | | | | | | | | | | | | | | | | | | | | | | |
DB 294 KSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 327

RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
AC P01870:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA CHAIN C REGION.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RL F-I haplotype."
RN [2] Immunogenetics 18:387-397(1983).
```

```
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RL immunoglobulin G of different allotype.";
RN [3] Biochem. J. 151:337-349(1975).
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RL heavy chain and identification of two genomic C gamma genes.";
RN [4] Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RL immunoglobulin G.";
RN [5] Biochem. J. 116:249-259(1970).
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
Stockholm (1967).
CC -! MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M16426; AAA31289.1; -.
DR PIR: A02161; GHRB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; Igc1; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 58.7%; Score 910.5; DB 1; Length 323;
Best Local Similarity 52.7%; Pred. No. 9.6e-63;
Matches 178; Conservative 31; Mismatches 54; Indels 75; Gaps 6;

QY 2 QAPGT---APCSRGWSADLCKMDCASCRRPHSDFCIGCAAP--PAPFLLWRS-- 54
||| | | | | | | | | | | | | | | | | | | | | | | | |
DB 4 KAPSVFPLAPC-----CGDTPSPSTVLGCLVKGYLPEPVTVTVNSGT 45
```



```
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00407; Igc1. 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
FT SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;
SQ
Query Match 54.0%; Score 838; DB 1; Length 326;
Best Local Similarity 60.5%; Pred. No. 3.4e-57;
Matches 150; Conservative 43; Mismatches 47; Indels 8; Gaps 3;
QY 40 CAAAPPAPE-----BLLWRSCDKTHTCPAPAEAGAPSVFLFPKPKDTLMISRTPEV 94
DB 82 CNVAHPASTKVKDKIVPNCG--GDCKPCICTGSE--VSSVFIFPPKPKDVITILTTPKV 138
QY 95 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREQYNSTYRVSVLTVFLHODMVLNGREY 154
DB 139 TCVVVDISQDDPEVHFVFNVDVDEVHTAQTTPPEQFNSTPERSVSELPILHODMVLNGRTF 198
QY 155 KCKVSNKALPAPIETISKAKQPREPOVYITLPPSRREMTKNQVSLTCLVKGFFYPSDTAV 214
DB 199 RCKVTSAAFPSPFIETISKPEGRTPQVHYTNMSPTEKTEMTQNEVSTICVMKGFYPPDIYV 258
QY 215 EWESNGQPENNYKTPPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCVSVYHAEALHNHYTQK 274
DB 259 ENQMNGQEQYKNTPTPTMDTDCGSLYSLKLVNKKVKEWQQGNFTFCVSLHGLHNHHTK 318
QY 275 SLSLSPGK 282
DB 319 SLSHSPGK 326
RESULT 8
GC3_MOUSE
ID GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; J00451; -: NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00407; Igc1. 2.
DR SMART; SM00410; Ig-like; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT SEQUENCE 329 AA; 36228 MW; F4582717A182BAD6 CRC64;
SQ
Query Match 53.9%; Score 836; DB 1; Length 329;
Best Local Similarity 62.7%; Pred. No. 4.9e-57;
Matches 156; Conservative 33; Mismatches 54; Indels 6; Gaps 2;
QY 40 CAAAPPAPEPRLWRSCD----KTHTCP--PCPAPEAGAPSVFLFPKPKDTLMISRTPE 93
DB 81 CNVAHPASKTELKRIEPRIPKPTPPGSCCPGNGILGSPSVFIETPKPKDALMISLTPE 140
QY 94 VTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREQYNSTYRVSVLTVFLHODMVLNGRE 153
DB 141 VTCVVVDVSEDDPDVHVSFVNDKVEHTAQTTPPEAQYNSTFRVVSALPIQHDWMRGKE 200
QY 154 YCKVSNKALPAPIETISKAKQPREPOVYITLPPSRREMTKNQVSLTCLVKGFFYPSDIA 213
DB 201 FKCKVNNKALPAPIETISKPKGKRAQTPOVYITPPPREQMSKKVSLTCLVTFNFFSEAS 260
QY 214 VWESNGQPENNYKTPPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCVSVYHAEALHNHYTQ 273
DB 261 VEWERNGELEQYKNTPTPTILDSDGYFLYSLKLTVDTSLQGEIFTCVSVYHAEALHNHTQ 320
QY 274 KSLSLSPGK 282
DB 321 KNLSRSPGK 329
RESULT 9
GC1_MOUSE
ID GC1_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
RT cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
```

103
RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salsar W.;
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
RT heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=78242288; PubMed=98524;
RA Adetugbo K.;
RT "Evolution of immunoglobulin subclasses. Primary structure of a
RT murine myeloma gammal chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
RN [5]
RP DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).
RN
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -;
DR EMBL; V00793; CAA24173.1; -;
DR EMBL; V00793; CAA24174.1; -;
DR EMBL; V00793; CAA24175.1; -;
DR EMBL; V00795; CAA24176.1; -;
DR PIR; A02159; GIMS.
DR GlycosuitedB; P01868; -;
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igc1; 2.
DR POSITIVE; PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON-TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 N-LINKED (GLCNAC. . .).
FT CARBOHYD 174 174 /FTID-CAR_000055.
FT
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 53.7%; Score 833; DB 1; Length 324;
Best Local Similarity 55.3%; Pred. No. 8.le-57;
Matches 151; Conservative 48; Mismatches 60; Indels 14; Gaps 4;

QY 17 ADLKDCKMDCASCRRPFDCLGCAAPPAPE-----RLLRSCDKTKTCTPP--CPAPEA 69
Db 59 SDLYTLSSSVTVSPSPSTVTCNVAHPASSTVKDKIVPRDGG-----CKPCICTVPEV 114
QY 70 EGAPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREE 129

115 S---SVFIPPKPKDVLITLTTPKVTGVVDVDSKDDPEQVFSFVDDVEVHTAQTPREE 171
QY 130 QYNSTRVRSVLTVLHODWLNGLNGKEYCKVSNKALPAPIETISKAKGQPREQVYTLPPS 189
Db 172 QFNSTFRSVSELPIMHODWLNGLNGEFCRVNSAFAPIETISKTKGRKPAQVYTIPTPP 231
QY 190 REEMTKNOVSLTCLVKGFPYSDIAVWESNGOPENNYKTPPVLDSDGSPFLYSKLTVDK 249
Db 232 KEQMAKDKVSLTCTMTDFFPEDITVQWNGQPAENKTKPTQIMNTNGSFYVSKLVNQK 291
QY 250 SRWQOQNVFSCSYNHEALHNHYTKSLSPGK 282
Db 292 SNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 324

RESULT 10
GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
RT conserved transmembrane sequence and a 28-residue intracellular
RT domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RL immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RL immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```
CC -----
DR EMBL: V00793; CAA24172.1; -
DR EMBL: V00793; CAA24173.1; -
DR EMBL: V00793; CAA24174.1; -
DR PIR: B02159; GIMSI.
DR MGD: MGI:96446; Igh-4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 53.4%; Score 828; DB 1; Length 393;
Best Local Similarity 55.1%; Pred. No. 2.4e-56;
Matches 150; Conservative 48; Mismatches 60; Indels 14; Gaps 4;

QY 17 ADLRCMDCASRPHSDFCLGCAAPAPF-----RLLRSCDKHTFCPP--CPAPRA 69
DB 59 SLDYLTSSSVTPSPRSEITVCNVAHPASSTKYDKKIVPRDCG---CKPCICTVPEV 114
QY 70 EGAPSVFLPPPKDGLMISRTPEVTCVVVDVSHEDPEVKFNWYGVDSVVHNAKTPREE 129
DB 115 S---SVFIFFPPKPKDVLITLTPKVTVCVVVDISKDDPEVQFSWFDVDEVTQAQTPREE 171
QY 130 QYNSTYRVSVLTVLHODWLNCKEYKCKVSKNPKALPAIEKTIKAKGQPREQVYTLPPS 189
DB 172 QFNSTFRSVSELPIMHQDWLNKEPKRVNSAAPPAPIEKTIKTKGRKPAQVYTIPPP 231
QY 190 REEMTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK 249
DB 232 KQMAKDKVSLTCLMTDFPEDITVEWQWNGQPAENYKNTQIPMTNGSYFYVSKLNVQK 291
QY 250 SRWQOGNVFSCVMHEALHNHYTQKSLSLSPG 281
DB 292 SNWEAGNTFTCSVLHEGLHNHHTEKSLSHSPG 323

RESULT 11
GCB_RAT ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2B CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
CC "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RT
```

```
RL Gene 74:473-482(1988).
DR PIR: PS0018; PS0018.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; Igcl; 2.
DR SMART: SM00410; Iglike; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 53.2%; Score 825; DB 1; Length 333;
Best Local Similarity 64.0%; Pred. No. 3.4e-56;
Matches 146; Conservative 33; Mismatches 49; Indels 0; Gaps 0;

QY 55 CDKTHCTCPCPAPEAGAPSVFLFPKPKDGLMISRTPEVTCVVVDVSHEDPEVKENWYV 114
DB 106 CPTCTCHKCPVPELIGGSPVFIFFPKPKDILLISQNAKVTVCVVVDSEEPDVQFSWFV 165
QY 115 DGVENVHNAKTKPREQYNSTYRVSVLTVLHODWLNCKEYKCKVSKNPKALPAIEKTIK 174
DB 166 NNVEVHTAQTPREQYNSTFRVVSALPIQHDWMSGKEFKCKVNNKALPSPKTIKSP 225
QY 175 KGQPREQVYTLPPREEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLD 234
DB 226 KGLVRKQVYVVGVPTEQLTEQTSVLTCLTSGFLPNDIGVETNSGHIENKYNKTEPVMD 285
QY 235 SDGSFSLYKLVNDRKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 282
DB 286 SDGSFMYSKLVNDRSWDRAPFVCSVYVHEGLHNHVEKSIKSPGK 333

RESULT 12
GCB_MOUSE ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84014483; PubMed=6314258;
RA Konaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; J00451; AAB59655.1; -
DR PIR; A02155; G3MSM.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IG1; 2.
DR SMART; SM00410; IG1; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSHEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 53.2%; Score 825; DB 1; Length 398;
Best Local Similarity 62.3%; Pred. No. 4.2e-56;
Matches 154; Conservative 33; Mismatches 54; Indels 6; Gaps 2;

QY 40 CAAPAPAPFLLWMSCD-----KTHTCP--PCPAEAGAPSVFLFPKPKDITLMSRTP 93
DB 81 CNVAHPASKTELKRIEPRIPKPSPPGSCPGNILGSPSVFIFPKPKDALMSLTPK 140
QY 94 VTCVVVDVSHEDPPVKENWYVDGVEVINAKTPREOYNSTYRVSVLTVLHQDWLNGKE 153
DB 141 VTCVVVDVSEDDPPHVSFVNDKKEVTAWTPQREAGYNSTFRVVSALPIHQDWMRGKE 200
QY 154 YKCVSNKALPAPIEKTISKAKGPPQVYTLPPSREEMTKNOVSLTCLVKGYPSDIA 213
DB 201 FKCVNNKALPAPIERTISPKGRAQTPQVYTIIPPREQMSKKVSLTCLVTFNFFSAIS 260
QY 214 VEWESGQPNKYKTPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQ 273
DB 261 VEWERNGEQDYKNTPPILDSDGTLYSKLTVDTDSWLQGEIFTCSVVHEALHNHHTQ 320
QY 274 KSLSLP 280
DB 321 KNLRSRP 327

RESULT 13
GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2C CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
PP SEQUENCE FROM N.A.
RX MEDLINE=8816903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfré P., Waldmann H., Calabi F.;
RT *Sequence of a rat immunoglobulin gamma 2c heavy chain constant

RT region cDNA: extensive homology to mouse gamma 3.;
RL Eur. J. Immunol. 18:317-319(1988).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X07189; CAA30169.1; -
DR PIR; S00847; S00847.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IG1; 2.
DR SMART; SM00410; IG1; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 52.7%; Score 817; DB 1; Length 329;
Best Local Similarity 53.8%; Pred. No. 1.4e-55;
Matches 155; Conservative 46; Mismatches 49; Indels 38; Gaps 3;

QY 6 TAPCSRGSWS-----ADLDKCMDCASCRRPHSDFLGCAAPAPFRLWRS 54
DB 69 TVPSSTWSTQVTCVSAHPATKSNLIKRIEPRPKRPPTDTC-----S 112
QY 55 CDKTHTCPPEAECAPSVFLFPKPKDITLMSRTPVTVVVDVSHEDPVEKENVY 114
DB 113 CD-----DNLGRPSVFIFPKPKDILMTLTLPKTVVVDVSEEDPVQFSWF 161
QY 115 DGVEVHNKTPREOYNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 174
DB 162 DNVRFVTAQTPHEEQNLNGTFRVSVTLRIHQDWMGSKGKCKVNNKDLPSPIEKTISK 221
QY 175 KGQPREPOVYTLPPSREEMTKNOVSLTCLVKGYPSDIAVWESNGQPNKYKTPPVLD 234
DB 222 RGRARTPQVYTIIPPREQMSKKVSLTCLVTFNFFSAISVSEWERNGEQDYKNTLPVLD 281
QY 235 SDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPCK 282
DB 282 SDESIFLYSKLTVDSWNRGDIYTCVHEALHNHHTQKSLSLSPCK 329

RESULT 14
GCC_RAT STANDARD; PRT; 322 AA.
AC P20760;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]

```
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 27 82
FT FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 136 196
FT FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9D0E1EDB CRC64;

Query Match 51.9%; Score 804.5; DB 1; Length 322;
Best Local Similarity 47.0%; Pred. No. 1.2e-54;
Matches 157; Conservative 38; Mismatches 62; Indels 77; Gaps 5;

QY 3 APTGAPCSRGSSWADLKDCMDCASCARPHSDFCIGCAAP--PAPFLLWRS----- 54
Db 12 AGTA-----LKSNWTLTGLVKGIFYPEPTVTWNSGALSSG 49
QY 55 -----CDKTH-----TCPPCPAPE 68
Db 50 VHTFPAVLQSGLYLTSSVTPSSWSSQAVTCNVAPASSPKVKIKVIPRECPCGCTG 109
QY 69 AGCAPSVLFPKPKDPTLMISRTPEVTCVVDVSHEDPEKFNWYVDGEVINATKPRE 128
Db 110 SE-VSSVFIFPPKTKDVLITLTPTKVTCTVVDVSDQNDPEVRFSWFIDVVEVHTAQT 168
QY 129 EQYNSTYRVSVLTVLHQDLNGLKGYCKVSNKALPAPIEKTISKAKGPQPPQVTLPP 188
Db 169 KQNSTLRSVSELPIVHRDLNGLTKFKCVNSGAPPAPIEKSISKEPTPGPQVYTPAP 228
QY 189 SREMTKNQVSLTCLVKGYPSPDIAVEWESNGQPNKYKTPPVLDSDGSEFFLYSKLTV 248
Db 229 PREEMTQSOVSITCMVKGYFPDPTDIETKMKNGQPQENYKNTPTMTDTSGLYSLKLVK 288
QY 249 KSRWQGNVFCVSMVEALHNHYTKSLSLSPGK 282
Db 289 KETWQGNFTTCSVLHGLHNHHTKSLSHSPGK 322

RESULT 15
ID GCAA_MOUSE STANDARD; PRT; 330 AA.
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RN Nucleic Acids Res. 8:3143-3155(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
RN Nucleic Acids Res. 9:1365-1381(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Olio R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
suggests that exons can be exchanged between genes in a multigenic
family.";
RN Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN [4]
RP MYELOMA PROTEIN MOPC 173
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
immunoglobulin: amino-acid sequence of the Fc fragment. Implications
for the evolution of immunoglobulin structure and function.";
RN Eur. J. Biochem. 43:423-435(1974).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1; -.
DR PIR; A02152; G2MSA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igc1; 2.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT FT DISULFID 144 204
FT FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 51.5%; Score 799.5; DB 1; Length 330;
Best Local Similarity 49.7%; Pred. No. 3e-54;
Matches 157; Conservative 29; Mismatches 69; Indels 61; Gaps 4;
```

```
QY 28 CRARPHSDFCLGCAAP--PAPFRLMRS----- 54
Db 15 CGDTTGSVTLGCLVKGYFFPEPVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTTSST 74
QY 55 -----CDKTH-----TCPP--CPAPEAGAPSVLFPKPKDYL 86
Db 75 WPSQITCNVAHPASSTKVDKIEPRGPTTKPCPPCKCAPNLLGGPSVFIFPKIKDVL 134
QY 87 MISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVWVSFLTVLHQ 146
Db 135 MISLPIVTCVVVDVSEDDPDVQISWFWNNVEVHTAQTQTHREDYNSTLRVVSALPIHQ 194
QY 147 DWLNGKEYCKKVNKALPAPIEKTISKAKGQPREPOVYITLPPSRREEMTKNOVSLTCLVKG 206
Db 195 DWMSGKEFKCKVNNKDLPAPIERTISKPGSVRAPOVYVLPPEEEMTKKQVTLTCNVTD 254
QY 207 FYPDSIAEVESNGQPENNYKTTPPVLDSDGSFELYKSLTVDKSRWOOQGNVFCSCVMHEA 266
Db 255 FMPEDYIYEWNTNGKTTELNTKTEPVLDSGYSFMYSKLRVEKKNWYERNYSYCSVYHEG 314
QY 267 LHNHYTOKLSLSLSPGK 282
Db 315 LHNHHTTKSFSPGK 330
```

Search completed: March 11, 2002, 15:23:52
Job time: 576 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:26 ; Search time 54.85 seconds
(without alignments)
752.029 Million cell updates/sec

Title: US-09-742-454A-7_COPY_28_309

Perfect score: 1551

Sequence: 1 EQAPGTAPCSRGSSWADLD.....MHEALNHHTQKSLSPGK 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	841	54.2	463	11 Q99LC4	Q99LC4 mus musculus
2	837	54.0	437	11 Q99LA4	Q99LA4 mus musculus
3	804	51.8	473	11 Q9D8L4	Q9D8L4 mus musculus
4	798.5	51.5	473	11 Q99L25	Q99L25 mus musculus
5	795.5	51.3	468	11 Q99L31	Q99L31 mus musculus
6	367	23.7	375	4 Q9BS21	Q9BS21 homo sapien
7	367	23.7	597	4 Q9BU10	Q9BU10 homo sapien
8	367	23.7	597	4 Q9BQ88	Q9BQ88 homo sapien
9	304.5	19.6	129	4 Q9NP84	Q9NP84 homo sapien
10	284.5	18.3	384	4 Q9UP60	Q9UP60 homo sapien
11	279.5	18.0	500	4 Q9BRV0	Q9BRV0 homo sapien
12	278.5	18.0	129	11 Q9QW3	Q9QW3 mus musculus
13	278.5	18.0	129	11 Q9CR75	Q9CR75 mus musculus
14	274.5	17.7	416	4 Q9NPP6	Q9NPP6 mus musculus
15	270	17.4	487	11 Q99KA4	Q99KA4 mus musculus
16	269.5	17.4	684	13 Q90544	Q90544 ginglymosto
17	269	17.3	426	11 Q9DCD9	Q9DCD9 mus musculus
18	255	16.4	479	11 Q99M22	Q99M22 mus musculus
19	255	16.4	484	11 Q99LA6	Q99LA6 mus musculus

20	185	11.9	268	13 Q90524	Q90524 ginglymosto
21	184.5	11.9	130	11 Q9D8W4	Q9D8W4 mus musculus
22	181	11.7	235	11 Q99M11	Q99M11 mus musculus
23	180	11.6	509	11 Q9QX57	Q9QX57 mus musculus
24	180	11.6	513	11 P97797	P97797 mus musculus
25	176.5	11.4	259	13 Q90530	Q90530 ginglymosto
26	174.5	11.3	509	11 Q9WTN4	Q9WTN4 mus musculus
27	173	11.2	509	11 Q98907	Q98907 mus musculus
28	171.5	11.1	257	13 Q90536	Q90536 ginglymosto
29	170	11.0	237	13 Q90545	Q90545 ginglymosto
30	168	10.8	105	11 Q99JC1	Q99JC1 mus musculus
31	168	10.8	506	6 Q46632	Q46632 bos taurus
32	167.5	10.8	252	13 Q90568	Q90568 ginglymosto
33	166.5	10.7	267	13 Q90529	Q90529 ginglymosto
34	165	10.6	238	7 Q9MXA2	Q9MXA2 aulonocara
35	165	10.6	506	6 Q46631	Q46631 bos taurus
36	162	10.4	261	7 Q19363	Q19363 sus scrofa
37	160	10.3	261	7 Q62868	Q62868 sus scrofa
38	157.5	10.2	237	7 Q9MX99	Q9MX99 aulonocara
39	155	10.0	261	7 Q98263	Q98263 sus scrofa
40	152.5	9.8	354	4 Q9NQK8	Q9NQK8 homo sapien
41	150	9.7	208	7 Q9MXA0	Q9MXA0 aulonocara
42	150	9.7	260	7 P79551	P79551 homo sapien
43	149.5	9.6	354	4 Q9P1W8	Q9P1W8 homo sapien
44	149	9.6	238	11 Q99M37	Q99M37 mus musculus
45	147.5	9.5	248	7 Q98044	Q98044 aulonocara

ALIGNMENTS

RESULT 1
Q99LC4 ID Q99LC4; PRELIMINARY; PRT; 463 AA.
AC Q99LC4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AAH03435.1; -
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match	54.2%	Score	841;	DB	11;	Length	463;
Best Local Similarity	60.0%	Pred. NO.	9.5e-68;				
Matches	150;	Conservative	42;	Mismatches	44;	Indels	14;
Gaps	4;						
QY	40	CAAAAPAPF-----RLWRSCKDTHTCPP--CPAPEAGCAPSVLEFPKPKOTLMISRTP	92				
DB	221	CNVAHPASSTKYDKKIVPRDCG-----CKPCICTVPEVS---SVFIFPPKPKDVLITLTP	273				
QY	93	EYTCVVVDVSHEDPEVKFNWYVDGVEFHNKTKPREEQYNSTYRVVSVLTVLDHQLMNGK	152				
DB	274	KVTCVVVDISKDDPEVQFSWFDVDDVEVHTAQTQPREQENSTFSVSELPIMHQDLNKG	333				
QY	153	EYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDI	212				
DB	334	EFKCRVNSAAFPAPIEKTISKTKGKPAQVVTIPPPKQMAKDKVSLTCLMTIDFPEDI	393				
QY	213	AVENESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSMHEALHNHT	272				
DB	394	TVEWQNGQPAENYKNTQPIDTDSGTFTYISKLVNQKSNWEAGNTFTCSVLHGLNNHT	453				
QY	273	OKSLSLSPGK 282					

```
Db 454 EKLSHSPGK 463
RESULT 2
Q9R1A4 PRELIMINARY; PRT; 437 AA.
ID Q9R1A4
AC Q9R1A4;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AF152372; AAD40243.1; -.
DR MGD; MGI:1924014; 1810060009rik.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IGL; 2.
DR SMART; SM00407; IGL; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IGLike; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR NON_TER 1 437
DR NON_TER 437 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 54.0%; Score 837; DB 11; Length 437;
Best Local Similarity 59.6%; Pred. No. 2e-67;
Matches 149; Conservative 43; Mismatches 44; Indels 14; Gaps 4;
QY 40 CAAPAPPP-----RLWRSCKDTHTCPP--CPAPEAGAPSVFLFPKPKDTLMISRT 92
Db 195 CNVAHPASSTKVDKVIIPDCG-----CKPCICTVPEVS---SVFIPPKPKDVLITLTP 247
QY 93 EYTCVVYDVSHEDPEVKFNKYVDGVEVHNNAKTRPREQYNSTYRVSVLVFLVHODWLNK 152
Db 248 KVTCCVVYDIDSKDDEVOQFSWFVDVVEHTAQTPREQFNSTFRSVSELPIMHQDWLNK 307
QY 153 EYCKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFPSDI 212
Db 308 EFKCRVNSAFAPIEKTISKGRKAPQVYIIPPKQKQAKDKVSLTCTMTDFFPEDI 367
QY 213 AVESNGSQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRWQQGVNFCVSNVHEALNHYT 272
Db 368 TVEQWNGQPAENYKNTQPTMDTDGSGYFVYSKLVNQRKSNWEAGNTFTCSVHEGLNHH 427
QY 273 OKLSLSLSPGK 282
Db 428 EKNLSHSPGK 437

RESULT 3
Q9D8L4 PRELIMINARY; PRT; 473 AA.
ID Q9D8L4
AC Q9D8L4;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE 1810060009RIK PROTEIN.
GN 1810060009RIK.
OS Mus musculus (Mouse).

Query Match 51.8%; Score 804; DB 11; Length 473;
Best Local Similarity 58.1%; Pred. No. 2.1e-64;
Matches 147; Conservative 38; Mismatches 66; Indels 2; Gaps 1;
QY 30 ARPHSDPCLGCAAPAPFPRLLRSCDKTHTCPPAPEAGAPSVFLFPKPKDTLMIS 89
Db 223 AHPASSTKVDKKTIEPRVP--ITQNPCLPEKCPCAAPDLGGPSVFIPPKIKDVLMS 280
QY 90 RTEVTCVVYDVSHEDPEVKFNKYVDGVEVHNNAKTRPREQYNSTYRVSVLVFLVHODWL 149
Db 281 LSPMVTCCVVYDVSHEDPDVQISWFFNVNVEHTAQTPHREDYNTSLRWVSALPIHQDWM 340
QY 150 NGREYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFY 209
Db 341 SGREFKCKVNNRALPSPIEKTISKGRPVAPQVYIIPPKQKQAKDKVSLTCTMTGFLP 400
QY 210 SDIAVWESNGQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRWQQGVNFCVSNVHEALN 269
Db 401 AEIADVDTWSNGRTEQYKNTATVLDSDGSGYFMYSKLVQRKSTWERSLGFACSVVHEGLN 460
QY 270 HYTKLSLSLSPGK 282
Db 461 HLTKTITSRLGK 473

RESULT 4
Q99L25 PRELIMINARY; PRT; 473 AA.
ID Q99L25
```

```
AC Q99L25;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AAH0388.1; -.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B7986DA155 CRC64;

Query Match 51.5%; Score 798.5; DB 11; Length 473;
Best Local Similarity 45.7%; Pred. No. 6.6e-64;
Matches 164; Conservative 30; Mismatches 80; Indels 85; Gaps 6;

QY 9 CSRGSS-----WSADLDCMDCAS-----CDKTH----- 59
Db 115 CSRGSIYGYGLYFDYWGQCTTIVSSAKTTAPSVPLAPVCGDGTGGSTVTLGCLVRG 174
QY 45 --PAPFRLIWR-----TCPP--CPAPEAGAPSVFLFPPKPKDTL 59
Db 175 YPEPVTLTWNSGSLSSGVHTTTPAVLQSDLYTLSSSVTVTSWTSQSTCNVAHPASST 234
QY 60 -----TCPP--CPAPEAGAPSVFLFPPKPKDTLMSRTPEVTCTVVDVSH 103
Db 235 KVDDKIEPRGPTIKPCPPCKPAPNLGGPSVFIIPPKIKDVLMSLSPMVTCVVVDVSE 294
QY 104 EPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODMLNGKEYCKVSNKAL 163
Db 295 DDPDQVQISWFNVNVEVLTQOTHTREDYNTSLRVVSALPIQHDWMSGEKFKVNNKAL 354
QY 164 PAPIEKTISKAGQREPOVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE 223
Db 355 PAPIERTISKAGSVRAPOVYVLPPEEEMTKKQVTLTQCMVDFDPEDYIVEMTNGKTE 414
QY 224 NNYKTPPVLDSDGSFFLYSKLTVDSKRWQGNVFSQSVHAEALHNHYTKLSLSPGK 282
Db 415 LNYKTEPVLDSDGSFYFYSKLRVEKKNWERNYSQSVHAEGLHNHHTTKFSRTPGK 473

RESULT 5
Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE SIMILAR TO RIKEN CDNA 1810060009 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AAH0388.1; -.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B332ADB CRC64;

Query Match 51.3%; Score 795.5; DB 11; Length 468;
Best Local Similarity 49.7%; Pred. No. 1.2e-63;
Matches 157; Conservative 29; Mismatches 69; Indels 61; Gaps 4;

QY 28 CRARPHSDFCLGCAAP--PAPFRLIWR----- 54
Db 153 CGDTTSSVTLGCLVKGFYFPEVPTLTWNSGSLSSGVHTTTPAVLQSDLYTLSSSVTVTSST 212
```

```
QY 55 -----CDKTH-----TCPP--CPAPEAGAPSVFLFPPKPKDTL 86
Db 213 WFSQSTCNVAHPASSTKVDDKIEPRGPTIKPCPPCKPAPNLGGPSVFIIPPKIKDVL 272
QY 87 MISRTPEVTCTVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHQ 146
Db 273 MISLSPMVTCVVVDVSEDDPDVQISWFNVNVEVLTQOTHTREDYNTSLRVVSALPIQHQ 332
QY 147 DWLNGKEYCKVSNKALPAPIEKTISKAGQREPOVYTLPPSREMTKNQVSLTCLVKG 206
Db 333 DWMSGKEFKCKVNNKALPAPIERTISKPGSVRAPOVYVLPPEEEMTKKQVTLTQCMVTD 392
QY 207 FYPSDIAVEWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDSKRWQGNVFSQSVHAE 266
Db 393 FMPEDIYVEMTNGKTELKNTPEVLDSDGSFYFYSKLRVEKKNWERNYSQSVVHGE 452
QY 267 LHNHYTKLSLSPGK 282
Db 453 LHNHHTTKFSRTPGK 468

RESULT 6
Q9BSZ1 PRELIMINARY; PRT; 375 AA.
AC Q9BSZ1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:10455).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004476; AAH04476.1; -.
SQ SEQUENCE 375 AA; 41314 MW; BLA0A098F473619 CRC64;

Query Match 23.7%; Score 367; DB 4; Length 375;
Best Local Similarity 33.8%; Pred. No. 3.8e-25;
Matches 75; Conservative 50; Mismatches 89; Indels 8; Gaps 6;

QY 67 PAEAGAPSVFLFPPKPKDTLMSRTPEVTCTVVDVSHEDPEVKFNMYVDGVEVHNNAKTP 126
Db 138 PDDTAIRVAIPPS-FASIFLTKSTKLCVLTDLTYD-SVTISWTRONGEAVKTHNI 195
QY 127 REQYNSTYRVSVLTVLHODMLNGKEYCKVSNKALPAPIEKTISKAGQREPOVYT 185
Db 196 SESHNPATFSAVGEASICEDDWNSEGRFTCTVTHTDLPSPKQITSRPKGVALLHRPDYL 255
QY 186 LPPSREMT-KNOVSLTCLVKGFYPSDIAVEWESNGOP--ENNYKTPPVLD--SDGSFF 240
Db 256 LPPEAQLNLRESATITCLVTGSPADVFYQNMQRGQPLSPKQITSRPKGVALLHRPDYL 315
QY 241 LYSKLTVDKSRWQGNVFSQSVHAEALHNHYTKLSLSPGK 282
Db 316 AHSILVSEEWNTGETYTCVVAHEALPNRVTERTVDKSTGK 357

RESULT 7
Q9BU10 PRELIMINARY; PRT; 597 AA.
AC Q9BU10;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:1652).
OS Homo sapiens (Human).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC002963; AAH02963.1; 2DFA8FB7E055851 CRC64;
SQ SEQUENCE 597 AA; 65274 MW; 2DFA8FB7E055851 CRC64;

Query Match 23.7%; Score 367; DB 4; Length 597;
Best Local Similarity 33.8%; Pred. No. 6.5e-25;
Matches 75; Conservative 50; Mismatches 89; Indels 8; Gaps 6;

QY 67 PEAGGARSVFLPPKPKDLMISTPEVTCVVDVSHEDPEVKFNWYDVGVEVINATKTP 126
   | : | | | : : : : : : : : : : : : | : | : |
Db 360 PQDQTAIRVFAIPPS-FASIFLTAKSTKLTCVLTDLTYD-SVTISWTRQNGEAVKTHNI 417
   | : | | | : : : : : : : : : : : : | : | : |
QY 127 REEQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIEKTISKAKGP-REPOVYT 185
   | : | | | : : : : : : : : : : : : | : | : |
Db 418 SESHPNATFFSANGVASCIEDDWSNGERTCTVTHDLPSPKQITSRPKGVALLHRPDVYL 477
   | : | | | : : : : : : : : : : : : | : | : |
QY 186 LPPSREMT-KNOVSLTCLVKGFYPSDIAVWESNGQP--ENNYKTTTPPVLD--SDGSFF 240
   | : | | | : : : : : : : : : : : : | : | : |
Db 478 LPPAREQLNLRRESATITCLVTGFSPADVFQVMQRGOPLSPEKYVTSAPMPEQAPGRYF 537
   | : | | | : : : : : : : : : : : : | : | : |
QY 241 LYSKLTVDKSRWQGNVFSVMEALHNHYTKLSLSLSPGK 282
   | : | | | : : : : : : : : : : : : | : | : |
Db 538 AHSILTVSEEWNTGETYTCVVAHEALPNRVTVTVDKSTGK 579
   | : | | | : : : : : : : : : : : : | : | : |

RESULT 8
Q980B8 ID Q9BQ88 PRELIMINARY; PRT; 597 AA.
AC Q9BQ88
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE UNKNOWN (PROTEIN FOR MGC:1905) (PROTEIN FOR MGC:1228).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC006180; AAH06180.1;
DR EMBL: BC001872; AAH01872.1;
SQ SEQUENCE 597 AA; 65300 MW; 2DAFAD50A6375851 CRC64;

Query Match 23.7%; Score 367; DB 4; Length 597;
Best Local Similarity 33.8%; Pred. No. 6.5e-25;
Matches 75; Conservative 50; Mismatches 89; Indels 8; Gaps 6;

QY 67 PEAGGARSVFLPPKPKDLMISTPEVTCVVDVSHEDPEVKFNWYDVGVEVINATKTP 126
   | : | | | : : : : : : : : : : : : | : | : |
Db 360 PQDQTAIRVFAIPPS-FASIFLTAKSTKLTCVLTDLTYD-SVTISWTRQNGEAVKTHNI 417
   | : | | | : : : : : : : : : : : : | : | : |
QY 127 REEQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIEKTISKAKGP-REPOVYT 185
   | : | | | : : : : : : : : : : : : | : | : |
Db 418 SESHPNATFFSANGVASCIEDDWSNGERTCTVTHDLPSPKQITSRPKGVALLHRPDVYL 477
   | : | | | : : : : : : : : : : : : | : | : |
QY 186 LPPSREMT-KNOVSLTCLVKGFYPSDIAVWESNGQP--ENNYKTTTPPVLD--SDGSFF 240
   | : | | | : : : : : : : : : : : : | : | : |

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB035480; BAA94792.1;
DR EMBL: AF191148; AAF69108.1;
DR EMBL: BC002718; AAH02718.1;
KW Signal; Transmembrane.
FT SIGNAL 9 31 POTENTIAL.
SQ SEQUENCE 129 AA; 13911 MW; BF3FDFB9C1E1C448 CRC64;

Query Match 19.6%; Score 304.5; DB 4; Length 129;
Best Local Similarity 60.0%; Pred. No. 4.8e-20;
Matches 60; Conservative 2; Mismatches 15; Indels 23; Gaps 2;

QY 1 EQAPGTAPCSRGSWSADLDKCMDCASCARPHSDFCGLGCAAAPAPPRLL----- 51
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 28 EQAPGTAPCSRGSWSADLDKCMDCASCARPHSDFCGLGCAAAPAPPRLLWPIILGALS 87
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 52 -----WRSCDKTHT-CPPCPAPEAGAPSVFL 77
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 88 LTFVLGLLGLVWRRRCRRREKFTTPIETGEGCPAVAL 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
Q9UP60 ID Q9UP60 PRELIMINARY; PRT; 384 AA.
AC Q9UP60
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SNC73 PROTEIN.
GN SNC73.
OS Homo sapiens (Human).
```

```
Db 478 LPPAREQLNLRRESATITCLVTGFSPADVFQVMQRGOPLSPEKYVTSAPMPEQAPGRYF 537
   | | | | | : : : | | | | | | | | | | | | | | | | | | | |
QY 241 LYSKLTVDKSRWQGNVFSVMEALHNHYTKLSLSLSPGK 282
   | : | | | : : : : : : : : : : : : | : | : |
Db 538 AHSILTVSEEWNTGETYTCVVAHEALPNRVTVTVDKSTGK 579
   | : | | | : : : : : : : : : : : : | : | : |

RESULT 9
Q9NP84 ID Q9NP84 PRELIMINARY; PRT; 129 AA.
AC Q9NP84
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TYPE I TRANSMEMBRANE PROTEIN PRECURSOR (TYPE I TRANSMEMBRANE PROTEIN
   FNI4).
GN FNI4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Tanaka S., Sugimachi K.;
RT "Human homologue of Fni4.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=20216634; PubMed=10751351;
RA Feng S.-L.Y., Guo Y., Factor V.M., Thorgerirsson S.S., Bell D.W.,
   Testa J.R., Peifley K.A., Winkles J.A.;
RT "The Fni4 immediate-early response gene is induced during liver
   regeneration and highly expressed in both human and murine
   hepatocellular carcinoma.";
RL Am. J. Pathol. 156:1253-1261(2000).
RN [3]
RC SEQUENCE FROM N.A.
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB035480; BAA94792.1;
DR EMBL: AF191148; AAF69108.1;
DR EMBL: BC002718; AAH02718.1;
KW Signal; Transmembrane.
FT SIGNAL 9 31 POTENTIAL.
SQ SEQUENCE 129 AA; 13911 MW; BF3FDFB9C1E1C448 CRC64;
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zheng S., Cao J., Cao W., Cai X., Geng L.;
 RT "Identification and characterization of SNCT73, a gene which is down-
 regulated in colorectal cancer."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
 CC DOMAIN.
 DR EMBL: AF067420; AAC19365.1; -.
 DR HSSP: P01825; 7FAB.
 DR InterPro: IPR003597; Ig_c1.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_3.
 DR SMART: SM00407; IgC1; 2.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;
 Query Match 18.3%; Score 284.5; DB 4; Length 384;
 Best Local Similarity 26.0%; Pred. No. 1.1e-17;
 Matches 89; Conservative 38; Mismatches 138; Indels 77; Gaps 12;
 QY 13 SWSADLDKCMDCASCARPHSDFCIGCAAA---PPAPFLLWRS----- 54
 DB 30 SSASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQELPSVTWSESGQGVTAARNFPSPD 89
 QY 55 -----CDKTHTC-----PPCPAPEAGCAPSVFLFPPK 81
 DB 90 ASGDLTYTSSQLTLPATQCLACKSVCHVKHYTNPSQDVTVPCTPSTPTSPST--PT 148
 QY 82 PK-----DTLMTISRTPEVTCVVVDVSHEDPEVKFNKALPAPIEKTISKAKQPREQVYTL 126
 DB 149 PSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGL-RDASGVFTTWPSSGK--SAVQGP 205
 QY 127 REEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKQPREQVYTL 186
 DB 206 PERDLGCGYSVSVLPGCAEPNHNKGTFTCTAAYPESKPTPLTATLSKS-GNTRFPEVHLL 264
 QY 187 PPSREEMTKNQ-VSLTCLVKGYFSPDIAVEWESNGO--PENNYKTTTPPVLD--SDG--SFF 240
 DB 265 PPPSEELALNELVTLTCLARGSPKDVLRWLQSGQELPREKYLTVASRQEPSQGTITFA 324
 QY 241 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 282
 DB 325 VTSILRVAEDWKKGDTFCMVGHEALPLAFTQKTIDRLAGK 366
 RESULT 11
 Q9BRV0
 ID Q9BRV0 PRELIMINARY; PRT; 500 AA.
 AC Q9BRV0;
 DT 01-JUN-2001 (TremBLrel. 17, Created)
 DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
 DE UNKNOWN (PROTEIN FOR MGC:14586).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC005951; AAH05951.1; -.
 SQ SEQUENCE 500 AA; 54154 MW; 0A9BF43F2A3CC6D9 CRC64;
 Query Match 18.0%; Score 279.5; DB 4; Length 500;
 Best Local Similarity 25.4%; Pred. No. 4.1e-17;

Matches 87; Conservative 41; Mismatches 137; Indels 77; Gaps 12;
 QY 13 SWSADLDKCMDCASCARPHSDFCIGCAAA---PPAPFLLWRS----- 54
 DB 146 SSASPTSPKVFPLSLCSTQPDGNVVIACLVQGFPPQELPSVTWSESGQGVTAARNFPSPD 205
 QY 55 -----CDKTHTC-----PPCPAPEAGCAPSVFLFPPK 81
 DB 206 ASGDLTYTSSQLTLPATQCLACKSVCHVKHYTNPSQDVTVPCTPSTPTSPST--PT 264
 QY 82 PK-----DTLMTISRTPEVTCVVVDVSHEDPEVKFNKALPAPIEKTISKAKQPREQVYTL 126
 DB 265 PSPSCCHPRLSLHRPALEDLLLGSEANLTCTLTGL-RDASGVFTTWPSSGK--SAVQGP 321
 QY 127 REEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKQPREQVYTL 186
 DB 322 PERDLGCGYSVSVLPGCAEPNHNKGTFTCTAAYPESKPTPLTATLSKS-GNTRFPEVHLL 380
 QY 187 PPSREEMTKNQ-VSLTCLVKGYFSPDIAVEWESNGO--PENNYKTTTPPVLD--SDG--SFF 240
 DB 381 PPPSEELALNELVTLTCLARGSPKDVLRWLQSGQELPREKYLTVASRQEPSQGTITFA 440
 QY 241 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 282
 DB 441 VTSILRVAEDWKKGDTFCMVGHEALPLAFTQKTIDRLAGK 482
 RESULT 12
 Q9QZW3
 ID Q9QZW3 PRELIMINARY; PRT; 129 AA.
 AC Q9QZW3;
 DT 01-MAY-2000 (TremBLrel. 13, Created)
 DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
 DE TYPE I TRANSMEMBRANE PROTEIN FN14.
 GN FGFRP2 OR FN14.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=20020297; PubMed=10551889;
 RA Meighan-Mantha R.L., Hsu D.K.W., Guo Y., Brown S.A.N., Feng S.-L.Y.,
 RA Peifley K.A., Alberts G.F., Copeland N.G., Gilbert D.J., Jenkins N.A.,
 RA Richards C.M., Winkles J.A.;
 RT "The mitogen-inducible Fn14 gene encodes a type I transmembrane
 protein that modulates fibroblast adhesion and migration."
 RL J. Biol. Chem. 274:33166-33176(1999).
 DR EMBL: AF156164; AAF07882.1; -.
 DR MGD; MGI:1351484; Fgfrp2.
 KW Transmembrane.
 SQ SEQUENCE 129 AA; 13637 MW; 14B5C68EEF493385 CRC64;
 Query Match 18.0%; Score 278.5; DB 11; Length 129;
 Best Local Similarity 56.0%; Pred. No. 1.1e-17;
 Matches 56; Conservative 2; Mismatches 19; Indels 23; Gaps 2;
 QY 1 EQAPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCIGCAAAAPPAPFRL----- 51
 DB 28 EQAPGTAPCSRGSSWSADLDKCMDCASCARPHSDFCIGCAAAAPPAPFRLWPILGALS 87
 QY 52 -----WRSCDKTHT--CPFCPAPEAGCAPSVFL 77
 DB 88 LVLVLAVSSFLWRRRCRRREKFTTTEETGGEGCGVAL 127
 RESULT 13
 Q9CR75
 ID Q9CR75 PRELIMINARY; PRT; 129 AA.
 AC Q9CR75;

```
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FIBROBLAST GROWTH FACTOR REGULATED PROTEIN 2.
GN FGFRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA, AND CEREBELLUM;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Rono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK005530; BAB24101.1; -
DR EMBL; AK005382; BAB23989.1; -
DR MGD; MGI:1351484; Fgfrp2.
SQ SEQUENCE 129 AA; 13641 MW; 1665C68B4D9A9253 CRC64;

Query Match 18.08; Score 278.5; DB 11; Length 129;
Best Local Similarity 56.08; Pred. No. 1.1e-17;
Matches 56; Conservative 2; Mismatches 19; Indels 23; Gaps 2;

Qy 1 EQAPGTAPCSRGSSWSADLDCMDCASCARPHSDFCILGCAAAAPPAPFRLL----- 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 28 EQAPGTSPCSSGSSWSADLDCMDCASCARPHSDFCILGCAAAAPPAPFRLLWPILGGALS 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 52 -----WRSCDKTHT-CPPCPAPAEAGAPSVFL 77
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 88 LVLVLALVSPFLVWRRCRRRREKFTTPIETGGEGCGPGVAL 127
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansoorge W., Ballabio A., Estivill X., Gibson K.,
RA Leirach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated molecular analysis of
```

```
RT human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN
DR EMBL; AL389978; CAB97534.1; -
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 17.78; Score 274.5; DB 4; Length 416;
Best Local Similarity 29.68; Pred. No. 9.3e-17;
Matches 81; Conservative 38; Mismatches 116; Indels 39; Gaps 12;

Qy 21 KCMD-CASCAR----PHSDFCLGCAAAAPPAPFRLLWRSCDKTHTCPPCPAPAEAGAPS 74
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 152 QCPGKSVTCHVKHYTNPSQDVTYPCVPPP-----PPCCHPR----- 189
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 75 VFLPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTRPEQYINST 134
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 190 LSLHRPALED-LLLGSEANLTCTLTGL-RDASGATFTWTPSSGK--SAVQGPPELDLGGC 245
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 135 YRVSVLTVLHQLDNLNGKEYCKVSNKALPAPIETKISKAKGQPREPOVYITLPPSRERM 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 246 YSVSVLPFGCAQPNHGETFTCTAAHELPKLTPLTANITKS-GNTPRPEVHLLPPSEELA 304
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 195 KNO-VSLTCLVKGYFSPDAVWEESNGO--PENNYKTTTPVLD-SDG--SFFLYSKLTVD 248
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 305 LNELVTLCCLARGSPKDVLRWLGSGSELPREKLTWASRQEPSQGITFAVTSILRVA 364
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 249 KSRMQQGNVFCVSMHEALHNYHTQKLSLSPGK 282
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 365 AEDWKKGDTFCMVGHEALPLAFTQKTDRLAGK 398
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
Q99KA4 PRELIMINARY; PRT; 487 AA.
AC Q99KA4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:6727).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004786; AAH04786.1; -
SQ SEQUENCE 487 AA; 52554 MW; 7DC8E96DB333077B CRC64;

Query Match 17.48; Score 270; DB 11; Length 487;
Best Local Similarity 30.68; Pred. No. 2.8e-16;
Matches 77; Conservative 41; Mismatches 100; Indels 34; Gaps 11;

Qy 40 CAAAPPAPFRLLWRSCDKTHTCPPCPAPAEAGAPSVFLPPPKPKDTLMISRTPEVTCVV 99
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 243 CSGPPP-----PCPCP-PSCH--PSLSLQRPALD-LLLGSDASLTCTLN 284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 100 DVSHEDPEVKFNKYVDGVVHNNAKTRPEQYINST---YRVSVLTVLHQLDNLNGKEYCK 156
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 285 GLRNPAGAV-FTW-----EPSTGKDAVQKAVQNSCGYSSVSLPGCAERNNSGASPK 338
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:14:13 ; Search time 53.92 Seconds
(without alignments)
71.436 Million cell updates/sec

Title: US-09-742-454a-7_COPY_28_79
Perfect score: 301
Sequence: 1 EQAPGTAPCSRGSNSADLD.....HSDFCIGCAAPAPFRLWL 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	301	100.0	114	20 AAW73409 Human secreted pro
2	301	100.0	129	20 AAW86506 Human liver clone
3	301	100.0	129	21 AAU57940 Human transmembran
4	301	100.0	129	22 AAU03498 Human TWEAK recept
5	301	100.0	309	22 AAU03500 Human TWEAK recept
6	147.5	49.0	112	21 AAY91463 Human secreted pro
7	147.5	49.0	155	21 AAY91604 Human secreted pro
8	147.5	49.0	156	21 AAY91552 Human secreted pro
9	67.5	22.4	242	22 AAG92250 C glutamicum prote
10	67	22.3	928	17 AAR97853 Rat REK7 eph-relat
11	67	22.3	1005	20 AAW83147 Rat receptor tyros

12	65.5	21.8	187	20	AAV29189	Amino acid sequenc
13	64	21.3	77	20	AAV12612	Human 5' EST secre
14	64	21.3	115	20	AAV59693	Secreted protein 4
15	64	21.3	115	21	AAB08524	Protein encoded by
16	64	21.3	115	21	AAV65397	Human 5' EST relat
17	64	21.3	115	22	AAG93299	Human protein HP10
18	64	21.3	120	21	AAV53361	Human colon cancer
19	64	21.3	141	20	AAV73883	Human prostate tum
20	64	21.3	141	20	AAV59770	Human normal ovari
21	64	21.3	478	22	AAB48033	Egf receptor relat
22	62.5	20.8	405	19	AAW33737	Epidermal growth f
23	62.5	20.8	509	14	AAR38210	LD2D3D4 EGF recept
24	62.5	20.8	529	14	AAR38209	LD1D2D3 Apal EGF r
25	62.5	20.8	621	21	AAV67309	Epidermal growth f
26	62.5	20.8	1210	21	AAV19259	Amino acid sequenc
27	62.5	20.8	1210	21	AAV50616	Human EGF receptor
28	62.5	20.8	1210	22	AAB68420	Amino acid sequenc
29	62.5	20.8	1251	16	AAV79475	Mouse LTPB-3. Mus
30	62.5	20.8	1251	22	AAB61481	Murine LTPB-3 prot
31	62.5	20.8	1251	22	AAB61483	Human TANGO 300 ex
32	62.5	20.8	3075	19	AAW50892	Human laminin A ch
33	62	20.6	234	22	AAW24485	Human EST encoded
34	62	20.6	567	22	AAV70531	Human PRO1 protein
35	62	20.6	720	21	AAV88280	Human TANGO 215 pr
36	62	20.6	720	21	AAV66695	Membrane-bound pro
37	62	20.6	720	22	AAU00401	Human secreted pro
38	62	20.6	720	22	AAV87544	Human PRO1344. HO
39	62	20.6	720	22	AAV70532	Human PRO2 protein
40	62	20.6	720	22	AAB65218	Human PRO1344 (UNQ
41	62	20.6	737	22	AAB93670	Human protein sequ
42	61	20.3	392	21	AAB38394	Human secreted pro
43	60.5	20.1	271	21	AAV19706	Protease-resistant
44	60.5	20.1	271	21	AAV19707	Protease-resistant
45	60.5	20.1	271	22	AAB68046	Amino acid sequenc

ALIGNMENTS

RESULT	1
AAW73409	
ID	AAW73409 standard; Protein; 114 AA.
AC	AAW73409;
XX	
DT	19-FEB-1999 (first entry)
XX	
DE	Human secreted protein encoded by Gene No. 13.
XX	
KW	Secreted protein; human; protein therapy; gene therapy; blood disorder;
KW	pathological condition; diagnosis; cancer; neurological disorder; disease;
KW	developmental abnormality; foetal deficiency; leukaemia; hepatic disease;
KW	immune system disorder; Alzheimer's disease; cognitive disorder;
KW	schizophrenia; prostate disease; autoimmune disorder; AIDS.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 114
FT	/note= "unspecified amino acid"
XX	
PN	WO9854206-A1.
XX	
PD	03-DEC-1998.
XX	
PF	28-MAY-1998; 98WO-US10868.
XX	
PR	29-AUG-1997; 97US-0056296.
PR	30-MAY-1997; 97US-0044039.
PR	30-MAY-1997; 97US-0048093.
PR	30-MAY-1997; 97US-0048101.
PR	30-MAY-1997; 97US-0048190.
PR	30-MAY-1997; 97US-0048356.

PR 30-MAY-1997: 97US-0050935.
PR 29-AUG-1997: 97US-0056250.
PR 29-AUG-1997: 97US-0056293.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Carter KC, Dillon PJ, Endress GA, Feng P, Ni J;
PI Rosen CA, Ruben SM, Yu G;
XX WPI; 1999-070209/06.
DR N-PSDB; AAV08823.
XX
XX New isolated human genes - useful for diagnosis and treatment of,
PT e.g. cancers, neurological disorders, immune diseases, developmental
PT disorders or blood disorders
XX
XX Claim 11; Page 153; 188pp; English.
PS
XX This sequence is encoded by a cDNA of the invention, designated
CC Gene No.13. This sequence represents a human secreted protein, and is
CC expressed in keratinocytes and to a lesser extent in endothelial
CC cells and placenta.
CC The DNA sequences of the invention and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the DNA sequences.
CC Specific uses are described for each of the DNA sequences and the encoded
CC proteins, based on which tissues they are most highly expressed in, and
CC include developing products for the diagnosis or treatment of cancer,
CC tumours, neurological disorders, developmental abnormalities and foetal
CC deficiencies, blood disorders, leukaemias, diseases of the immune system
CC (including allergies or asthma), hepatic disease, Alzheimer's and
CC cognitive disorders, schizophrenia, prostate diseases, autoimmune
CC disorders and AIDS. The polypeptides are also useful for identifying
CC their binding partners.
XX
XX Sequence 114 AA;
SQ
Query Match 100.0%; Score 301; DB 20; Length 114;
Best Local Similarity 100.0%; Pred. No. 3e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EQAPGTAPCSGRGSSWSADLDKCMDCASCARPHSDFCGCAAAPAPFRLLW 52
Db 28 eqapgtapcsgrgsswsadldkcmdcascrarphsdflcgcaappapfrllw 79
RESULT 2
AAW88506
ID AAW88506 standard; Protein; 129 AA.
XX
XX AAW88506;
AC AAW88506;
XX
XX 30-MAR-1999 (first entry)
DT
XX
XX Human liver clone HP10432-encoded membrane protein.
DE
XX
XX Transmembrane protein; HP10432; human; liver.
KW
XX
XX Homo sapiens.
OS
XX
XX WO9855508-A2.
PN
XX
XX 10-DEC-1998.
PD
XX
XX 03-JUN-1998; 98WO-JP02445.
PF
XX
XX 03-JUN-1997; 97JP-0144948.
PR
XX
XX (PROT-) PROTEGENE INC.
PA
XX (SAGA) SAGAMI CHEM RES CENTRE.

XX Kato S, Sekine S, Yamaguchi T;
PI
XX WPI; 1999-045730/04.
DR N-PSDB; AAV84374.
XX
XX New human proteins containing transmembrane domains and their
PT encoding sequences - useful in the preparation of antibodies and
PT large-scale protein production, gene diagnosis, and gene therapy
XX
XX Claim 1; Page 152-153; 178pp; English.
PS
XX This is the amino acid sequence of a transmembrane protein encoded
CC by human liver cDNA clone HP10432 (see AAV84374). The encoded protein
CC has a signal-like N-terminal region and one internal transmembrane
CC domain. The invention provides nucleotide sequences (see AAV84359-76)
CC coding for 18 transmembrane proteins (see AAV88491-508), vectors
CC containing such polynucleotides, and eukaryotic cells containing the
CC vectors. The proteins can be used as antigens or as compositions
CC in the preparation of antibodies against the proteins. The
CC polynucleotides can be used as probes for gene diagnosis, and as
CC gene sources for gene therapy and large-scale production of proteins
CC encoded by the cDNA. The host cells are used for the detection of
CC ligands corresponding to the expressed proteins, and the screening
CC of low mol.wt. medicines.
XX
XX Sequence 129 AA;
SQ
Query Match 100.0%; Score 301; DB 20; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.4e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EQAPGTAPCSGRGSSWSADLDKCMDCASCARPHSDFCGCAAAPAPFRLLW 52
Db 28 eqapgtapcsgrgsswsadldkcmdcascrarphsdflcgcaappapfrllw 79
RESULT 3
AAW57940
ID AAY57940 standard; Protein; 129 AA.
XX
XX AAY57940;
AC AAY57940;
XX
XX 23-MAR-2000 (first entry)
DT
XX
XX Human transmembrane protein HTMPN-64.
DE
XX
XX Human; transmembrane protein; HTMPN; diagnosis; immunospecific;
KW antiproliferative; neuroprotective; immune disorder;
KW reproductive disorder; smooth muscle disorder; neurological disorder;
KW gastrointestinal disorder; developmental disorder;
KW cell proliferative disorder.
XX
XX Homo sapiens.
OS
XX
XX WO961471-A2.
PN
XX
XX 02-DEC-1999.
PD
XX
XX 28-MAY-1999; 99WO-US11904.
PF
XX
XX 29-MAY-1998; 98US-0087260.
PR
XX 02-JUL-1998; 98US-0091674.
PR
XX 02-OCT-1998; 98US-0102954.
PR
XX 24-NOV-1998; 98US-0109869.
PR
XX (INCY-) INCYTE PHARM INC.
PA
XX
XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;
PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;
PI Au-Young J;
XX

FT	Protein	/note= "Signal peptide"	
FT		28..309	
FT		/note= "Mature human TWEAKR-Fc fusion protein.	
FT		Specifically referred to in Claim 11"	
FT	Domain	28..79	
FT		/note= "From TWEAKR extracellular domain. Specifically	
FT		referred to in Claims 4 and 10"	
FT	Region	80..81	
FT		/note= "From a BglII cloning site"	
FT	Region	82..309	
FT		/note= "Fc portion"	
XX	WO200145730-A2.		
PN			
XX	28-JUN-2001.		
XX			
XX	19-DEC-2000; 2000WO-US34755.		
XX			
XX	20-DEC-1999; 99US-0172878.		
PR	10-MAY-2000; 2000US-0203347.		
XX			
XX	(IMV) IMMUNEX CORP.		
PA			
XX	Wiley SR;		
PI			
XX	WPI; 2001-417975/44.		
XX	N-PSDB; AAS03965.		
XX			
PT	Modulating angiogenesis in a mammal for treating diseases mediated by		
PT	angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or		
PT	peripheral tissue, by administering antagonist or agonist of TWEAK		
PT	receptor		
XX			
XX	Claim 4; Page 45-46; 46pp; English.		
XX			
CC	The sequence represents a fusion protein consisting of the human TWEAK		
CC	receptor (TWEAKR) protein extracellular domain fused to an Fc portion		
CC	from human IgG1. This fusion protein, TWEAKR-Fc, is used in the		
CC	preparation of TWEAKR agonists and antagonists. The TWEAK protein is a		
CC	member of the tumour necrosis factor (TNF) family and induces		
CC	angiogenesis. TWEAKR may therefore be used to screen for and develop		
CC	TWEAKR agonists and antagonists for the modulation of angiogenesis, to be		
CC	used in the treatment and diagnosis of human disease. The disorders		
CC	mediated by angiogenesis include ocular disorders characterised by ocular		
CC	neovascularisation such as diabetic retinopathy, neovascular glaucoma,		
CC	retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,		
CC	rubeosis, uveitis, macular degeneration and corneal graft		
CC	neovascularisation, and inflammatory diseases such as arthritis,		
CC	rheumatism and psoriasis. Other treatable diseases include malignant and		
CC	metastatic conditions such as sarcomas and carcinomas, benign tumours and		
CC	preneoplastic conditions, myocardial angiogenesis, haemophilic joints,		
CC	scleroderma, vascular adhesions, atherosclerotic plaque		
CC	neovascularisation, telangiectasia, wound granulation, coronary		
CC	atherosclerosis, peripheral atherosclerosis and ischaemia.		
XX			
SQ	Sequence 309 AA;		

DT	29-JUN-2000	(first entry)	
XX			
DE	Human secreted protein sequence encoded by gene 13 SEQ ID NO:136.		
XX			
KW	Human; secreted protein; diagnosis; cytostatic; immunosuppressive;		
KW	antiHIV; antiinflammatory; neutropic; neuroprotective; antiallergic;		
KW	osteoarthritic; antiarthritic; antibacterial; antidiabetic; antiasthma;		
KW	antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;		
KW	immune disease; inflammation; blood disorder; tumour; chromosome 16.		
OS	Homo sapiens.		
XX			
FN	WO2000006698-A1.		
XX			
PD	10-FEB-2000.		
XX			
PF	29-JUL-1999; 99WO-US17130.		
XX			
PR	30-JUL-1998; 98US-0094657.		
PR	05-AUG-1998; 98US-0095486.		
PR	06-AUG-1998; 98US-0095454.		
PR	06-AUG-1998; 98US-0095455.		
PR	12-AUG-1998; 98US-0096319.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;		
PI	Latleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;		
PI	Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;		
XX			
XX			
DR	WPI: 2000-195282/17.		
DR	N-PSDB: AAA26358.		
XX			
PT	New isolated human genes and the secreted polypeptides they encode,		
PT	useful for diagnosis and treatment of e.g. cancers, neurological		
PT	disorders, immune diseases, inflammation or blood disorders -		
XX			
XX			
PS	Claim 11; Page 465; 634pp; English.		
XX			
CC	The polynucleotide sequences given in AAA26346 to AAA26458 encode the		
CC	human secreted proteins given in AAY91451 to AAY91691. The human secreted		
CC	proteins can have activities based on the tissues and cells they are		
CC	expressed in. Examples of the activities are: cytostatic;		
CC	immunosuppressive; antiHIV; antiinflammatory; neutropic; neuroprotective;		
CC	antiallergic; osteoarthritic; antiarthritic; antibacterial; antidiabetic;		
CC	antiasthma; antipsoriatic; and cardiant. The polynucleotides and their		
CC	corresponding secreted proteins are useful for preventing, treating or		
CC	ameliorating medical conditions, e.g. by protein or gene therapy. Also		
CC	pathological conditions can be diagnosed by determining the amount of the		
CC	proteins in a sample or by determining the presence of mutations in the		
CC	polynucleotides. Specific uses are described for each of the		
CC	polynucleotides, based on which tissues they are most highly expressed		
CC	in, and include developing products for the diagnosis or treatment of		
CC	cancer, tumours, neurodegenerative disorders, developmental abnormalities		
CC	and foetal deficiencies, blood disorders, diseases of the immune system,		
CC	autoimmune diseases, hepatic and renal disease, inflammation,		
CC	allergies, Alzheimer's and behavioural disorders, schizophrenia,		
CC	osteoporosis, arthritis, infections, AIDS, spinal cord injuries,		
CC	transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,		
CC	cardiovascular disorders, reproductive disorders, gastrointestinal		
CC	disorders, respiratory disorders and metabolic disorders. The proteins		
CC	or polynucleotides can also be used as food additives or preservatives.		
CC	The proteins are also useful for identifying their binding partners.		
CC	AAA26337 to AAA26345 and AAY91450 are sequences used in the		
CC	exemplification of the present invention.		
XX			
SQ	Sequence 112 AA;		
XX			

QY 1 EQAPGTAPCGSSWSADLDKCMDCa-SC 28
CC |||||||||||||||||||f: ||
Db 28 eqapgtapcrgsswsadldkcmdcstsc 56

RESULT 7
AAAY91604
ID AAY91604 standard; Protein; 155 AA.
XX
AC AAY91604;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 13 SEQ ID NO:277.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nontropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour; chromosome 16.
XX
OS Homo sapiens.
XX
PN WO200006698-A1.
XX
PD 10-FEB-2000.
XX
PF 29-JUL-1999; 99WO-US17130.
XX
PR 30-JUL-1998; 98US-0094657.
PR 05-AUG-1998; 98US-0095486.
PR 06-AUG-1998; 98US-0095454.
PR 06-AUG-1998; 98US-0095455.
PR 12-AUG-1998; 98US-0096319.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
WPI; 2000-195282/17.
XX
New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX
Disclosure: Page 36-37; 634pp; English.
XX
The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nontropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.

CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 155 AA;

Query Match 49.0%; Score 147.5; DB 21; Length 155;
Best Local Similarity 93.1%; Pred. No. 4.2e-09;
Matches 27; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 EQAPGTAPCGSSWSADLDKCMDCa-SC 28
CC |||||||||||||||||||f: ||
Db 28 eqapgtapcrgsswsadldkcmdcstsc 56

RESULT 8
AAAY91552
ID AAY91552 standard; Protein; 156 AA.
XX
AC AAY91552;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 13 SEQ ID NO:225.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nontropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
PN WO200006698-A1.
XX
PD 10-FEB-2000.
XX
PF 29-JUL-1999; 99WO-US17130.
XX
PR 30-JUL-1998; 98US-0094657.
PR 05-AUG-1998; 98US-0095486.
PR 06-AUG-1998; 98US-0095454.
PR 06-AUG-1998; 98US-0095455.
PR 12-AUG-1998; 98US-0096319.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
WPI; 2000-195282/17.
XX
N-PSDB; AAA26447.
XX
New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX
Claim 11; Page 528; 634pp; English.
XX
The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nontropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed

CC (AAT18897).
XX
SQ Sequence 928 AA;

Query Match 22.3%; Score 67; DB 17; Length 928;
Best Local Similarity 31.6%; Pred. No. 20;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;
Qy 4 PGTAPCSGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAP 45
Db 273 ppkmhcsaegwlvpgkcmckagyeekngtcqvcvrfkaspshstcskc---pp 326

RESULT 11
AAW83147
ID AAW83147 standard; Protein; 1005 AA.
XX
AC AAW83147;
XX
DT 11-FEB-1999 (first entry)
XX
DE Rat receptor tyrosine kinase Etk-1.
XX
KW Receptor tyrosine kinase; Ror-1; Ror-2; Etk-1; Etk-2; detection;
KW neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
KW binding protein; BDNF; NT-3; diagnosis.
XX
OS Rattus sp.
XX
PN US5843749-A.
XX
PD 01-DEC-1998.
XX
PF 06-JUN-1995; 95US-0469537.
XX
PR 17-MAR-1995; 95US-0406247.
PR 26-JUL-1991; 91US-0736559.
PR 28-OCT-1993; 93US-0144992.
PR 06-JUN-1995; 95US-0469537.
XX
PA (REG-) REGENERON PHARM INC.
XX
PI Maisompierre PC, Masiakowski P, Yancopoulos GD;
XX
DR WPI: 1999-044584/04.
DR N-PSDB: AAV70207.

XX
PT DNA encoding receptor tyrosine kinase proteins - and corresponding
PT proteins
XX
PS Example; Fig 22; 194pp; English.
XX
CC The present invention describes nucleic acid molecules for ror-1,
CC ror-2, etk-1 and etk-2. Also described are the corresponding proteins:
CC Ror-1; Ror-2; Etk-1; and Etk-2. The proteins are orphan receptor
CC tyrosine kinases. The present sequence represents rat Etk-1.
XX
SQ Sequence 1005 AA;

Query Match 22.3%; Score 67; DB 20; Length 1005;
Best Local Similarity 31.6%; Pred. No. 22;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;
Qy 4 PGTAPCSGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAP 45
Db 273 ppkmhcsaegwlvpgkcmckagyeekngtcqvcvrfkaspshstcskc---pp 326

RESULT 12
AAW29189
ID AAW29189 standard; Protein; 187 AA.

XX AAY29189;
AC
XX
DT 25-OCT-1999 (first entry)
XX

XX Amino acid sequence of a virulence factor encoded by ORF24368.
DE
XX
KW Human pathogen; virulence polypeptide; virulence factor;
KW pathogenic infection; Pseudomonas aeruginosa infection.
XX
OS Pseudomonas aeruginosa.
XX
PN W09927129-A1.
XX
PD 03-JUN-1999.
XX
PF 25-NOV-1998; 98WO-US25247.
XX
PR 25-NOV-1997; 97US-0066517.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;
PI Rahme LG, Tan M, Tsongalis J;
XX
DR WPI: 1999-357851/30.
XX
PT Virulence factors useful in developing disease treatments
XX
PS Disclosure; Fig 3; 228pp; English.

XX The present sequence represents a Pseudomonas aeruginosa polypeptide
CC sequence. P. aeruginosa is an opportunistic human pathogen present in
CC soil water and plants. The specification describes virulence polypeptides
CC and nucleic acid sequence encoding such polypeptides. These sequences
CC can be used to identify a compound which is capable of decreasing the
CC expression of a pathogenic virulence factor. Compounds that inhibit
CC the expression or activity of virulence factor polypeptides can be
CC used to treat pathogenic infections, especially where the infection
CC is a P. aeruginosa infection.
CC note: the sequences given in the specification were poorly legible, and
CC in some instances assumptions were made as to the identity of the
CC residue; it is therefore possible that the sequence given below is
CC not entirely correct.
XX
SQ Sequence 187 AA;

Query Match 21.8%; Score 65.5; DB 20; Length 187;
Best Local Similarity 39.5%; Pred. No. 6.2;
Matches 17; Conservative 2; Mismatches 11; Indels 13; Gaps 3;
Qy 6 TAPCSGSSWSADLDKCMDCASC-RARPHSDFCLGCAAAP 47
Db 123 sagcvpgpgwsg-----tgcapcsgrp-----caapppsp 153

RESULT 13
AAY12612
ID AAY12612 standard; Protein; 77 AA.
XX
AC AAY12612;
XX
DT 22-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID NO: 277 from WO 9906553.
XX
KW Human: secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.

FH Key Location/Qualifiers
FT Modified-site 83 /note= "potential phosphorylation site"
XX
PN WO200050588-A2.
PD 31-AUG-2000.
XX
PF 01-FEB-2000; 2000WO-US02595.
XX
PR 22-FEB-1999; 99US-0255381.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Walker MG, Volkmuth W, Klingler TM, Lal P;
XX
DR WPI; 2000-558397/51.
XX
PT Novel gene associated with colon cancer and coexpressed with one or
PT more known colon cancer genes in a number of biological samples, for
PT use in gene therapy
XX
PS Claim 4; Page 32-33; 33pp; English.
XX
CC The present sequence is encoded by a novel gene which is associated with
CC colon disease. The gene is coexpressed with one or more colon cancer
CC genes selected from carbonic anhydrase I, II and IV (CA I, II and IV),
CC carcinoembryonic antigen family of proteins (cea), colorectal carcinoma
CC tumour associated antigen (CO-029), down-regulated in adenoma (dra),
CC fatty-acid binding protein (fabp), galectin (galec), glutathione
CC peroxidase (gpx2), quanylin (guan), cytokeratin 8 and 20 (ker 8 and 20),
CC cadherin (cadher) or intestinal mucin (muc-2). The present polynucleotide
CC sequence is useful for diagnosing a disease or condition associated
CC with altered expression of the gene. The polynucleotide and encoded
CC polypeptides are useful for treating or preventing such diseases. They
CC are useful for diagnosing colon cancer, metastatic colon cancer, atrophic
CC gastritis, cholecystitis, Crohn's disease, irritable bowel syndrome and
CC ulcerative colitis.
XX
SQ Sequence 115 AA;

Query Match 21.3%; Score 64; DB 21; Length 115;
Best Local Similarity 30.8%; Pred. No. 5, 7;
Matches 16; Conservative 3; Mismatches 23; Indels 10; Gaps 2;
QY 2 QAP-----GTAPCSRGSSWADLDKCM-DCASCRARPHSDFCLGCAAA 43
||| | | : | : | | | | | | |
Db 4 qapvvvtqpgvgpapqnsnwtgmcdcfscgvcfcgfcgfcgvcgva 55

Search completed: March 11, 2002, 15:14:14
Job time: 98 sec

THIS PAGE BLANK (USPTO)

Db 273 PPKMCSAEGEWLVPICKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC-----PP 326

RESULT 2
US-08-440-815-2
; Sequence 2, Application US/08440815
; Patent No. 5798448
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; ATTORNEY/AGENT INFORMATION: AL-1 Neurotrophic Factor
; TITLE OF INVENTION: AL-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,815
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330128
; FILING DATE: 27-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 920C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/952-8674
; TELEFAX: 415/952-8674
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-440-815-2

Query Match 22.3%; Score 67; DB 1; Length 928;
Best Local Similarity 31.6%; Pred. No. 6.6;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAPP 45
Db 273 PPKMCSAEGEWLVPICKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC-----PP 326

RESULT 3
US-08-486-449-2
; Sequence 2, Application US/08486449
; Patent No. 6280732
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; ATTORNEY/AGENT INFORMATION: AL-1 Neurotrophic Factor
; TITLE OF INVENTION: AL-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,449
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330128
; FILING DATE: 27-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0920P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-486-449-2

Query Match 22.3%; Score 67; DB 4; Length 928;
Best Local Similarity 31.8%; Pred. No. 6.6;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAPP 45
Db 273 PPKMCSAEGEWLVPICKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC-----PP 326

RESULT 4
US-08-469-537A-103
; Sequence 103, Application US/08469537A
; Patent No. 5843749
; GENERAL INFORMATION:
; APPLICANT: Maisonnier, et al.
; TITLE OF INVENTION: EHK AND ROR TYROSINE
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,537A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/406,247
; FILING DATE: 17-MAR-1995
; APPLICATION NUMBER: USSN 08/144,992
; FILING DATE: 28-OCT-1993
; APPLICATION NUMBER: USSN 07/736,559
; FILING DATE: 26-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempner, Ph.D., Gail M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 1005 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-537A-103

Query Match 22.3%; Score 67; DB 2; Length 1005;
Best Local Similarity 31.6%; Pred. No. 7.1;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;
QY 4 PGTAFCSRGSSWSADLKCM-----DCASCR-----ARPHSDFCGCAAAP 45
DB 273 PPKMCSAEGEWLVPIGKCMKAGYEKNGTCQVCPGFFKASPHSQTCSC-PP 326

RESULT 5
US-08-336-708A-9
; Sequence 9, Application US/08336708A
; Patent No. 5521295
; GENERAL INFORMATION:
; APPLICANT: Pacifici, Robert E.
; APPLICANT: Thomason, Arlen R.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: Hybrid Receptor Molecules
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,708A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy
; REFERENCE/DOCKET NUMBER: A-241A
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-706A-9

Query Match 20.8%; Score 62.5; DB 1; Length 644;
Best Local Similarity 34.0%; Pred. No. 15;
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;
QY 9 CSRGSSWSADLKCM-----CAS-----CRARPHSDFCG-CAAAPAP 47
DB 194 CPNGSCWAGAGEENCQKLTKIIICAQCGRGCRGKSPSDCHNQCAGCTGP 243

RESULT 6
US-08-484-438-7
; Sequence 7, Application US/08484438
; Patent No. 5811098

Patent No. 5811098 5780031
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Cuifouscou, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstr m, Ingegerd
; APPLICANT: Hellstr m, Karl E.
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,438
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,442
; FILING DATE: 14-OCT-1994
; APPLICATION NUMBER: US 08/150,704
; FILING DATE: 10-NOV-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/981,165
; FILING DATE: 24-NOV-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-438-7

Query Match 20.8%; Score 62.5; DB 2; Length 1210;
Best Local Similarity 34.0%; Pred. No. 28;
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;
QY 9 CSRGSSWSADLKCM-----CAS-----CRARPHSDFCG-CAAAPAP 47
DB 194 CPNGSCWAGAGEENCQKLTKIIICAQCGRGCRGKSPSDCHNQCAGCTGP 243

RESULT 7
US-08-475-035-4
; Sequence 4, Application US/08475035
; Patent No. 5985553
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; APPLICANT: AARONSON, STUART A.
; TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE

/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NEEDLE & ROSENBERG, P.C.
/ STREET: Suite 1200, 127 Peachtree Street
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: USA
/ ZIP: 30303
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/475,035
/ FILING DATE: 7 Jun 1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Perryman, David G.
/ REGISTRATION NUMBER: 33,438
/ REFERENCE/DOCKET NUMBER: 1414.656
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404/688-0770
/ TELEFAX: 404/688-9880
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1210 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-475-035-4

Query Match 20.8%; Score 62.5; DB 2; Length 1210;
Best Local Similarity 34.0%; Pred. No. 28;
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;
QY 9 CSRGSSWSADLKDMD-----CAS-----CRARPHSDFCIG-CAAAAPPAP 47
Db 194 CPNGSCWAGGAEENCQKLTIKICACQCSGRCRGKSPSCCHNQCAAGCTGP 243

RESULT 8
PCT-US95-02251-3
/ Sequence 3, Application PC/TUS9502251
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
/ NUMBER OF SEQUENCES: 18
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Arnold, White & Durkee
/ STREET: P.O. Box 4433
/ CITY: Houston
/ STATE: Texas
/ COUNTRY: United States of America
/ ZIP: 77210
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
/ SOFTWARE: PatentIn Release #1.0, Version
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/02251
/ FILING DATE: CONCURRENTLY HERewith
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/316,650
/ FILING DATE: 30-SEP-1994
/ CLASSIFICATION:
/ APPLICATION NUMBER: US 08/199,780
/ FILING DATE: 18-FEB-1994

/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Parker, David L.
/ REGISTRATION NUMBER: 32,165
/ REFERENCE/DOCKET NUMBER: UMIC009p--
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (512) 418-3000
/ TELEFAX: (713) 789-2679
/ TELEX: 79-0924
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1251 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ PCT-US95-02251-3

Query Match 20.8%; Score 62.5; DB 5; Length 1251;
Best Local Similarity 29.4%; Pred. No. 29;
Matches 15; Conservative 9; Mismatches 12; Indels 15; Gaps 3;

QY 9 CSRGSSWS-----ADLDKCMDCASCRRPHSDFC-----LGCAAAPPAPF 48
Db 1000 CKQGFYDGNLLECDVDDECLDESNCR-----NGVCENTWRLPCACTPPAEY 1046

RESULT 9
US-08-199-780-3
/ Sequence 3, Application US/08199780
/ Patent No. 5763416
/ GENERAL INFORMATION:
/ APPLICANT: Bonadio, Jeffrey
/ TITLE OF INVENTION: Gene Transfer Into Bone Cells
/ TITLE OF INVENTION: And Tissues
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Arnold, White & Durkee
/ STREET: P.O. Box 4433
/ CITY: Houston
/ STATE: TX
/ COUNTRY: USA
/ ZIP: 77210
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30B
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/199,780
/ FILING DATE: 18-FEB-1994
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Parker, David
/ REGISTRATION NUMBER: 32,165
/ REFERENCE/DOCKET NUMBER: UMIC:002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (512) 320-7200
/ TELEFAX: (512) 474-7577
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1252 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-199-780-3

Query Match 20.8%; Score 62.5; DB 1; Length 1252;
Best Local Similarity 29.4%; Pred. No. 29;
Matches 15; Conservative 9; Mismatches 12; Indels 15; Gaps 3;

RESULT 11
US-08-460-309-5
; Sequence 5, Application US/08460309
; Patent No. 5837496
; GENERAL INFORMATION:
; APPLICANT: Engvall, Eva
; APPLICANT: Leivo, Ilmo
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
; TITLE OF INVENTION: Fragments and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
;

Patent No. 5872231
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/125,077
;; FILING DATE: 22-SEP-1993
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US PCT/US 94/10730
;; FILING DATE: 21-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/472,319
;; FILING DATE: 30-JAN-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/919,951
;; FILING DATE: 27-JUL-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Campbell, Cathryn A.
;; REGISTRATION NUMBER: 31,815
;; REFERENCE/DOCKET NUMBER: P-LA 9721
;; TELEPHONE: (619) 535-9001
;; TELEFAX: (619) 535-8949
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3075 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-125-077-5

Query Match 20.8%; Score 62.5; DB 2; Length 3075;
Best Local Similarity 23.5%; Pred. No. 73;
Matches 16; Conservative 6; Mismatches 23; Indels 23; Gaps 2;

QY 4 PGTAAPC-----SRGSSWSADLDKCMDCASCRCRAPHSDFC-----LGC 40
Db 1395 PLVAPCVPCSNHSDTCDPNTGKLCNGDNTAGDHCDVCTGYGKVTGSASDCALCAC 1454

QY 41 AAAPPAPF 48
Db 1455 PHSPPASF 1462

RESULT 13
US-08-976-255-11
;; Sequence 11, Application US/08976255
;; Patent No. 6136581
;; GENERAL INFORMATION:
;; APPLICANT: Jono, Keith E.
;; APPLICANT: Plowman, Gregory
;; TITLE OF INVENTION: KINASE GENES AND USES
;; NUMBER OF SEQUENCES: 53
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Lyon & Lyon
;; STREET: 633 West Fifth Street
;; STREET: Suite 4700
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90071-2066
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: IBM P.C. DOS 5.0
;; SOFTWARE: FastSeq for Windows 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/976,255
;; FILING DATE: NO. 6136581ember 21, 1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 60/031,675
;; FILING DATE: NO. 6136581ember 22, 1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warburg, Richard J.
;; REGISTRATION NUMBER: 32,327
;; REFERENCE/DOCKET NUMBER: 229/182
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1384 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: Protein
;; US-08-976-255-11

Query Match 20.1%; Score 60.5; DB 4; Length 1384;
Best Local Similarity 32.7%; Pred. No. 54;
Matches 16; Conservative 4; Mismatches 18; Indels 11; Gaps 2;

QY 4 PGTAAPC-----SRGSSWSADLDKCMDCASCRCRAPHSDFCGCAAPPA 46
Db 523 PGVVPVLSAHSPLSGSEYFIRLEAAPAG-----HDPDCAGCAPSPPA 566

RESULT 14
US-09-286-529-20
;; Sequence 20, Application US/09286529
;; Patent No. 6297367
;; GENERAL INFORMATION:
;; APPLICANT: Catherine Tribouley
;; TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES
;; FILE REFERENCE: 1408.003/200130.439C1
;; CURRENT APPLICATION NUMBER: US/09/286,529
;; CURRENT FILING DATE: 1999-04-05
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 20
;; LENGTH: 211
;; TYPE: PRT
;; ORGANISM: Homo sapien
;; US-09-286-529-20

Query Match 19.8%; Score 59.5; DB 4; Length 211;
Best Local Similarity 25.6%; Pred. No. 10;
Matches 21; Conservative 6; Mismatches 14; Indels 41; Gaps 5;

QY 4 PGT---APCSRGSS-----WSADLDKCMDC-----AS 27
Db 54 PGTFVQRPCCRRSDPTTCGCPCCPRHYTFQWY-LERCRCYNVLCGEREEARACHATHNRA 112
QY 28 CRAR-----PHSDFCGCAAPPA 45
Db 113 CRCRTGFFAHAGFCLEHASCPP 134

RESULT 15
US-09-286-529-17
;; Sequence 17, Application US/09286529
;; Patent No. 6297367
;; GENERAL INFORMATION:
;; APPLICANT: Catherine Tribouley
;; TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES
;; FILE REFERENCE: 1408.003/200130.439C1
;; CURRENT APPLICATION NUMBER: US/09/286,529
;; CURRENT FILING DATE: 1999-04-05
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 3.0


```
; SEQ ID NO 17
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-286-529-17
```

```
Query Match      19.8%; Score 59.5; DB 4; Length 299;
Best Local Similarity 25.6%; Pred. No. 14;
Matches 21; Conservative 6; Mismatches 14; Indels 41; Gaps 5;

QY  4 PGT---APCSRGS-----WSADLDKCMDC-----AS 27
    ||| ||| |
Db  54 PGTFFQPCRRDSTTCGPPRHYTQFWNY-LERCRCNVLCGEREEARACHATHNRA 112
    ||| ||| |
QY  28 CRAR---PHSDFCLGCAAPP 45
    ||| ||| |
Db  113 CRCRTGFFAHAGFCLEHASCPP 134
```

```
Search completed: March 11, 2002, 15:13:14
Job time: 38 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: March 11, 2002, 15:14:51 ; Search time 31.92 Seconds
(without alignments)
124.094 Million cell updates/sec
Title: US-09-742-454A-7_COPY_28_79
Perfect score: 301
Sequence: 1 EQAPGTAPCSRGSSWSADLD.....HSDFLGCAAPAPFRLIW 52
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	73.5	24.4	probable pyruvate
2	67	22.3	receptor-like tyro
3	67	22.3	receptor tyrosine
4	67	22.3	receptor-like tyro
5	67	22.3	receptor tyrosine
6	66	21.9	conserved hypotet
7	65.5	21.8	receptor-type prot
8	64	21.3	epidermal growth f
9	64	21.3	fibrillin-1 precu
10	62.5	20.8	epidermal growth f
11	62.5	20.8	latent transformin
12	62.5	20.8	laminin alpha-1 ch
13	62	20.6	notch 3 protein -
14	61	20.3	ferredoxin 2[4Fe-4
15	61	20.3	metallothionein 3
16	61	20.3	metallothionein 3,
17	61	20.3	growth inhibitory
18	61	20.3	hypothetical prote
19	61	20.3	pyruvate formate-1
20	60.5	20.1	TA20 protein - com
21	60.5	20.1	hypothetical prote
22	60.5	20.1	epidermal growth f
23	60.5	20.1	KIAA0641 protein -
24	60.5	20.1	low density lipopr
25	60	19.9	xanthine dehydrog
26	60	19.9	LDL receptor precu
27	60	19.9	serine proteinase
28	59.5	19.8	conserved hypotet
29	59	19.6	metallothionein 3,

variant-specific s
glypican precursor
receptor protein-t
protein-tyrosine k
metallothionein 3
hypothetical prote
hypothetical prote
regulatory protein
neurofibromatosis-
neurofibromin I -
neurofibromin I - mo
hemocytin - silkwo
chloromuconate cyc
hypothetical prote
factor H - bovine
receptor tyrosine

ALIGNMENTS
RESULT 1
G69099
probable pyruvate synthase (EC 1.2.7.1) gamma chain - Methanobacterium thermoautotrop
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-May-2000
C:Accession: G69099
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني,
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514
A:Accession: G69099
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-261 <MTH>
A:Cross-references: GB:AF0000929; GB:AF000666; NID:g2622853; PIDN:AA886210.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1740
C:Superfamily: pyruvate synthase gamma chain
C:Keywords: coenzyme A; oxidoreductase

Query Match 24.4%; Score 73.5; DB 2; Length 261;
Best Local Similarity 29.2%; Pred. No. 0.69;
Matches 19; Conservative 8; Mismatches 21; Indels 17; Gaps 4;
Qy 4 PGTAPCSRGSSWSA-----DLDKCMDCASCA-----RPHS---DFCLGCA-AAPPA 46
Db 191 PGSTVKNKTSWRTFKPVLDRKDKCIDCDNCILFCPEGCINREHEIDYDYCKGCGCAEKC 250
Qy 47 PFRLL 51
Db 251 PVKAI 255
RESULT 2
S51603
receptor-like tyrosine kinase Etk-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C:Accession: S51603
R:Maisonnier, P.C.; Barrezuela, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A:Title: Etk-1 and Etk-2: two novel members of the Eph receptor-like tyrosine kinase
A:Reference number: S49015; MUID:94067777
A:Accession: S51603
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-893 <MAI>
A:Cross-references: EMBL:S68028

A>Note: the authors translated the codon GAC for residue 170 as Glu
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat
C:Keywords: ATP; transmembrane protein
F:563-829/Domain: protein kinase homology <KIN>
F:571-579/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 893;
Best Local Similarity 31.6%; Pred. No. 8.4;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45
Db 273 PPKMHCSAEGEWLPVIGKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 326

RESULT 3

S47489
receptor tyrosine kinase - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S47489
R:Taylor, V.; Pfarr, S.; Miescher, G.C.; Honegger, P.; Breitschopf, H.; Lassmann, H.; St
submitted to the EMBL Data Library, April 1994
A:Description: Expression and developmental regulation of EHK-1, a neuronal ELK-like rec
A:Reference number: S47489
A:Accession: S47489
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-898 <TAY>
A:CROSS-references: EMBL:X78689; NID:g531543; PIDN:CAA55357.1; PID:g531544
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat
C:Keywords: ATP; transmembrane protein
F:568-834/Domain: protein kinase homology <KIN>
F:576-584/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 898;
Best Local Similarity 31.6%; Pred. No. 8.4;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45
Db 278 PPKMHCSAEGEWLPVIGKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 331

RESULT 4

S51604
receptor-like tyrosine kinase EHK-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 29-May-1998
C:Accession: S51604
R:Maisonpierre, P.C.; Barrezaeta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A:Title: EHK-1 and EHK-2: two novel members of the Eph receptor-like tyrosine kinase fam
A:Reference number: S49015; MUID:94067777
A:Accession: S51604
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-981 <MAI>
A:CROSS-references: EMBL:S68029
A>Note: the authors translated the codon GAC for residue 170 as Glu
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat
C:Keywords: ATP; transmembrane protein
F:651-917/Domain: protein kinase homology <KIN>
F:659-667/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 981;
Best Local Similarity 31.6%; Pred. No. 9;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45

Db 273 PPKMHCSAEGEWLPVIGKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 326

RESULT 5

S49015
receptor tyrosine kinase EHK-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C:Accession: S49015; S51602
R:Maisonpierre, P.C.; Barrezaeta, N.X.; Yancopoulos, G.D.
Oncogene 8, 3277-3288, 1993
A:Title: EHK-1 and EHK-2: two novel members of the Eph receptor-like tyrosine kinase
A:Reference number: S49015; MUID:94067777
A:Accession: S49015
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1005 <MAI>
A:CROSS-references: EMBL:S68024
A>Note: the authors translated the codon GAC for residue 170 as Glu
A:Accession: S51602
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-305, 'G', 359-1005 <MA2>
A:CROSS-references: EMBL:S68026
A>Note: the authors translated the codon GAC for residue 170 as Glu
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repea
C:Keywords: ATP; transmembrane protein
F:675-941/Domain: protein kinase homology <KIN>
F:683-691/Region: protein kinase ATP-binding motif

Query Match 22.3%; Score 67; DB 2; Length 1005;
Best Local Similarity 31.6%; Pred. No. 9.1;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAAPP 45
Db 273 PPKMHCSAEGEWLPVIGKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 326

RESULT 6

E83085
conserved hypothetical protein PA4488 [imported] - Pseudomonas aeruginosa (strain PAO
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83085
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.N.; Kas, A.; Larbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: E83085
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-549 <STO>
A:CROSS-references: GB:AE004862; GB:AE004091; NID:g9950716; PIDN:AAG07876.1; GSPDB:GN
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA4488
C:Superfamily: Escherichia coli hypothetical protein b2226

Query Match 21.9%; Score 66; DB 2; Length 549;
Best Local Similarity 36.7%; Pred. No. 7.5;
Matches 18; Conservative 3; Mismatches 16; Indels 12; Gaps 2;

QY 4 PGTAPCSRGSSWSADLDKCMDCASCRPHSDFCLGCAAAAPPFRLLW 52
Db 380 PASAPTRRTAAWSADL---VLGATVNVHSD-----QPGPARLSW 416

RESULT 7

I50615
receptor-type protein-tyrosine kinase Cek7, long splice form - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C:Accession: I50615; I50616; I50614
R:Stever, D.A.; Verderame, M.F.
Gene 148, 219-226, 1994
A:Title: Identification of a complete Cek7 receptor protein tyrosine kinase coding sequence
A:Reference number: I50614; MUID:95047429
A:Accession: I50615
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1013 <SIE>
A:Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60613.1; PID:g555618
A:Accession: I50616
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-572, 'R', 596-1013 <SI2>
A:Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60613.1; PID:g555619
A:Accession: I50614
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-279, 444-572, 'R', 596-1013 <SI3>
A:Cross-references: EMBL:U03910; NID:g555617; PIDN:AAB60614.1; PID:g555620
C:Genetics:
A:Gene: Cek7
C:Superfamily: protein-tyrosine kinase, receptor type eph; fibronectin type III repeat
C:Keywords: alternative splicing; Atp; transmembrane protein
F:649-915/Domain: protein kinase homology <KIN>
F:657-665/Region: protein kinase ATP-binding motif
F:938-1004/Domain: SAM homology <SAM>

Query Match 21.88; Score 65.5; DB 2; Length 1013;

Best Local Similarity 29.2%; Pred. No. 13;

Matches 19; Conservative 5; Mismatches 22; Indels 19; Gaps 4;

QY 1 EQAPTAPCRSGSSWSADLCKM-----DCASCR-----ARPHSDPCLCGAAAPP 45

Db 245 DEAP-KWHCSAEGEWLPIGKCLCKAGYEKNNTQCVCRPGFFKASHPSPCSKC---PP 300

QY 46 APFRL 50

: :

Db 301 HSYTL 305

RESULT 8

A53183
epidermal growth factor receptor precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.;
Genes Dev. 8, 399-413, 1994
A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
A:Reference number: A53183; MUID:94170986
A:Accession: A53183
A:Molecule type: mRNA
A:Residues: 1-1210 <LUE>
A:Cross-references: GB:U03425
R:Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
Oncogene 6, 673-676, 1991
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site
A:Reference number: A43818; MUID:91232866
A:Accession: A43818
A:Molecule type: mRNA
A:Residues: 1-714 <AVI>
A:Cross-references: GB:X59698
R:Eisinger, D.P.; Serrero, G.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24942
A:Accession: S24942

A:Molecule type: mRNA
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>
A:Cross-references: EMBL:Z12608
R:Heisermann, G.J.; Gill, G.N.
J. Biol. Chem. 263, 13152-13158, 1988
A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylate
A:Reference number: A28941; MUID:88330814
A:Accession: A28941
A:Molecule type: protein
A:Residues: 689-694, 'X', 696-704, 'L', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-10
R:Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.
submitted to the EMBL Data Library, April 1994
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Recept
A:Reference number: S45325
A:Accession: S45325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971, 'K', 973-1210 <VER>
A:Cross-references: EMBL:X78987; NID:9488830; PIDN:CAA55587.1; PID:9488831
R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Dev, S.K.
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
A:Title: Expression of the epidermal growth factor receptor gene is regulated in mous
A:Reference number: I49643; MUID:93126380
A:Accession: I49643
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 12-20, 22-132 <RES>
A:Cross-references: GB:I06864; NID:9193001; PIDN:AAA53029.1; PID:g567201
C:Genetics:
A:Gene: EGFR
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: Atp; growth factor receptor; kinase-related transforming protein; phospho
F:1-24/Domain: signal sequence #status predicted <SIG>
F:648-670/Domain: transmembrane #status predicted <TMMS>
F:712-977/Domain: protein kinase homology <KIN>
F:720-728/Region: protein kinase ATP-binding motif
F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental
F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental
F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 21.3%; Score 64; DB 2; Length 1210;
Best Local Similarity 31.7%; Pred. No. 22;
Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

QY 4 PGTAP-----CSRGSSWSADLCKMD-----CAS-----CRARPHSDPCLG-CAAAPAP 47

Db 184 PSSCPKCDPSCPNCGSGGGEQCQLTKIICAQCSHRCGRSPSCDCHNQCAAGCTGP 243

RESULT 9

A55624
fibrillin-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 11-Jan-2000
C:Accession: A55624
R:Yin, W.; Smiley, E.; Germiller, J.; Sanguinetti, C.; Lawton, T.; Pereira, L.; Ramire
J. Biol. Chem. 270, 1798-1806, 1995
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin
A:Reference number: A55624; MUID:95130561
A:Accession: A55624
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <YIN>
A:Cross-references: GB:L29454; NID:g575509; PIDN:AAA56840.1; PID:g575510
C:Genetics:
A:Gene: Fbn-1
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:1201-1236/Domain: EGF homology <EGF>

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:50 ; Search time 19.78 Seconds
(without alignments)
96.389 Million cell updates/sec

Title: US-09-742-454A-7_COPY_28_79

Perfect score: 301

Sequence: 1 EQAPCTAPCSRSSNSADLD.....HSDFCLGCAAAPAPFRLWL 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	73.5	24.4	81	1	PORD_METTH
2	67	22.3	1005	1	EPAS5_RAT
3	65.5	21.8	1013	1	EPAS5_CHICK
4	65	21.6	314	1	Y127_HUMAN
5	64	21.3	1210	1	EGFR_MOUSE
6	64	21.3	2871	1	FBN1_MOUSE
7	62.5	20.8	1210	1	EGFR_HUMAN
8	62.5	20.8	3075	1	LMAL_HUMAN
9	62	20.6	54	1	PER_PEPAS
10	62	20.6	2318	1	NTC3_MOUSE
11	61.5	20.4	495	1	MLP2_DRONE
12	61	20.3	66	1	MT3_RAT
13	61	20.3	68	1	MT3_MOUSE
14	60	19.9	837	1	LDLR_RABIT
15	60	19.9	1877	1	PKG5_MOUSE
16	59.5	19.8	431	1	YG84_METTH
17	59.5	19.8	547	1	Y213_HUMAN
18	59	19.6	68	1	Y213_HUMAN
19	59	19.6	558	1	GPCL_RAT
20	59	19.6	1037	1	EPAS5_HUMAN
21	58.5	19.4	1255	1	ERB2_HUMAN
22	58	19.3	68	1	MT3_BOVIN
23	58	19.3	422	1	Y140_HUMAN
24	58	19.3	443	1	FBL4_CRIGR
25	58	19.3	443	1	FBL4_MOUSE
26	57.5	19.1	624	1	NIFA_AZOLI
27	57.5	19.1	1696	1	PKL5_BRACL
28	57.5	19.1	2820	1	NFL_RAT
29	57.5	19.1	2839	1	NFL_HUMAN
30	57.5	19.1	2841	1	NFL_MOUSE
31	57.5	19.1	3133	1	HMCT_BOMMO
32	57	18.9	983	1	EPAS3_MOUSE
33	56.5	18.8	316	1	CC12_CABEL

34	56.5	18.8	316	1	CC13_CABEL
35	56	18.6	215	1	HP25_TANAS
36	56	18.6	684	1	FBL1_CHICK
37	56	18.6	1217	1	EGF_MOUSE
38	55.5	18.4	59	1	FER2_DESDN
39	55.5	18.4	279	1	SLI3_MOUSE
40	55.5	18.4	279	1	SLI3_RAT
41	55.5	18.4	1278	1	NPCI_HUMAN
42	55	18.3	252	1	DHSB_BACSU
43	55	18.3	394	1	FDHB_METTF
44	54.5	18.1	265	1	YIT6_YEAST
45	54.5	18.1	464	1	YMG8_YEAST

ALIGNMENTS

RESULT	PORD_METTH	1
ID	PORD_METTH	STANDARD; PRT; 81 AA.
AC	P56815;	2000 (Rel. 39, Created)
DT	30-MAY-2000	(Rel. 39, Last sequence update)
DT	30-MAY-2000	(Rel. 40, Last annotation update)
DE	20-AUG-2001	(Rel. 40, Last annotation update)
DE	PYRUVATE SYNTHASE SUBUNIT PORD (EC 1.2.7.1) (PYRUVATE OXIDOREDUCTASE DELTA CHAIN) (POR) (PYRUVIC-FERREDOXIN OXIDOREDUCTASE DELTA SUBUNIT).	
GN	PORD OR MTH1740.1.	
OS	Methanobacterium thermoautotrophicum.	
OC	Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;	
OC	Methanothermobacter.	
OX	NCBI_TaxID=145262;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=DELTA H;	
FX	MEDLINE=98037514; PubMed=9371463;	
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J., Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K., Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D., Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R., Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabakar S., McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M., Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.; "Complete genome sequence of Methanobacterium thermoautotrophicum delta: functional analysis and comparative genomics.";	
RL	J. Bacteriol. 179:7135-7155(1997).	
CC	-!- CATALYTIC ACTIVITY: PYRUVATE + COA + OXIDIZED FERREDOXIN = ACETYL-COA + CO(2) + REDUCED FERREDOXIN.	
CC	-!- COFACTOR: BINDS TWO 4FE-4S CLUSTERS.	
CC	-!- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE GAMMA CHAIN.	
CC	-!- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.	
CC	-!- CAUTION: THERE SEEMS TO BE A SEQUENCING ERROR THAT FUSES TOGETHER PORC AND PORD. WE HAVE CUT THE ORF INTO ITS TWO CONSTITUENTS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; AE000929; AAB86210.1; ALT_INIT.	
DR	HSP; P00195; ICLF.	
DR	InterPro; IPR001450; 4FE4S_ferdxin.	
DR	Pfam; PF00037; fer4; 2.	
DR	PRINTS; PR00353; 4FE4SPRDOXIN	
DR	PROSITE; PS00198; 4FE4S_FERREDOXIN; 2.	
KW	Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S; Complete proteome.	
KW	METAL 34	
FT	METAL 37	
FT	METAL 40	
FT	IRON-SULFUR 1 (4FE-4S) (POTENTIAL).	
FT	IRON-SULFUR 1 (4FE-4S) (POTENTIAL).	
FT	IRON-SULFUR 1 (4FE-4S) (POTENTIAL).	


```
Query Match      22.3%; Score 67; DB 1; Length 1005;
Best Local Similarity 31.6%; Pred. No. 3.2;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 4 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARHSDFCGLGCAAPP 45
DB 273 PPKMCSAEGEWLVPIGKCMKAGYEKNGTCVCRPGFFKASPHSQCSCRC---pp 326

RESULT 3
EPAS_CHICK STANDARD; PRT; 1013 AA.
AC P54755; 007495;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EPHRIN TYPE-A RECEPTOR 5 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE KINASE RECEPTOR CEK7).
GN EPHA5 OR CEK7.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body wall;
RX MEDLINE=95047429; PubMed=7958948;
RA Slevier D.A., Verderame M.F.;
RT "Identification of a complete Cek7 receptor protein tyrosine kinase
RT coding sequence and cDNAs of alternatively spliced transcripts.";
RL Gene 148:219-226(1994).
RN [2]
RP SEQUENCE OF 512-1013 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93288394; PubMed=8510926;
RA Sajjadi F.G., Pasquale E.B.;
RT "Five novel avian Eph-related tyrosine kinases are differentially
RT expressed.";
RL Oncogene 8:1807-1813(1993).
CC -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS TO
CC EPHRIN-A1, -A2, -A3, -A4 AND -A5. ABLE TO COLLAPSE GROWTH CONES.
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1, 2 AND 3 (SHOWN
CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: DETECTED IN THE 10-DAY EMBRYONIC BRAIN, WEAKER
CC EXPRESSION IN THE REST OF THE 10-DAY EMBRYO. UNDETECTED IN ADULT
CC TISSUES.
CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN
CC RECEPTOR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U03910; AAB60613.1; -
CC EMBL; U03910; AAB60614.1; -
CC EMBL; U03910; AAB60612.1; -
CC EMBL; Z19058; CAA79508.1; -
CC HSP; P00523; 2PTK
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001090; Ephrin_rcptor.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR001777; FN_III.
CC InterPro; IPR001426; Rptor_tyr_kin_v.
```

```
DR InterPro; IPR001660; SAM.
DR InterPro; IPR001245; Tyr_kin.
DR Pfam; PF01404; EPH_lbd; 1.
DR Pfam; PF00041; fn3_2
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD001495; Ephrin_rcptor; 1.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW Alternative splicing.
FT SIGNAL 1 31
FT CHAIN 32 1013
FT DOMAIN 32 549
FT TRANSMEM 550 570
FT DOMAIN 571 1013
FT DOMAIN 196 330
FT DOMAIN 331 440
FT DOMAIN 441 537
FT DOMAIN 651 912
FT DOMAIN 939 1013
FT SITE 1011 1013
FT NP_BIND 657 665
FT BINDING 683 683
FT ACT_SITE 776 776
FT MOD_RES 626 626
FT MOD_RES 632 632
FT MOD_RES 809 809
FT MOD_RES 958 958
FT CARBOHYD 240 240
FT CARBOHYD 275 275
FT CARBOHYD 345 345
FT CARBOHYD 399 399
FT CARBOHYD 412 412
FT CARBOHYD 437 437
FT VARSPPLIC 280 443
FT VARSPPLIC 573 595
FT CONFLICT 981 1013
FT SEQUENCE 1013 AA; 112245 MW; AC36FDAE8F38382 CRC64;
```

```
Query Match      21.8%; Score 65.5; DB 1; Length 1013;
Best Local Similarity 29.2%; Pred. No. 4.7;
Matches 19; Conservative 5; Mismatches 22; Indels 19; Gaps 4;

QY 1 EQAPGTCPSRGSSWSADLDKCM-----DCASCR-----ARHSDFCGLGCAAPP 45
DB 245 DEAP-KMCHSABGEWLVPIGKCLCKAGYEKNTQVCRPGFFKASPHSPSCSKC---pp 300

QY 46 APFRL 50
DB 301 HSYTL 305
```

```
RESULT 4
Y127_HUMAN STANDARD; PRT; 314 AA.
ID Y127_HUMAN
AC Q14140;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
```

20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN KIAA0127.
GN KIAA0127.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96127530; PubMed=8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:167-174(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).
CC -----
KW EMBL; D50917; BAA09476.1; -.
KW Hypothetical protein.
SQ SEQUENCE 314 AA; 33896 MW; E43107FC565AAC31 CRC64;

Query Match 21.6%; Score 65; DB 1; Length 314;
Best Local Similarity 33.3%; Pred. No. 2;
Matches 16; Conservative 6; Mismatches 24; Indels 2; Gaps 1;

QY 3 APTACRSGSSWSDLDKCMDCASCRCRPHSDFCLGCAAPAPPPRL 50
DB 112 SPSSHPCDLGS--TTPLACLTPASLLEDDDDTFC*SQAMQPTAPTKL 157

RESULT 5
EGFR_MOUSE STANDARD; PRT; 1210 AA.
AC Q01279;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112).
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=93026370; PubMed=1408137;
RA Avivi A., Skorecki K., Yayon A., Givol D.;
RT "Promoter region of the murine fibroblast growth factor receptor 2.
RT (bek/KGFR) gene.";
RL Oncogene 7:1957-1962(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C, AND CD-1; TISSUE=Liver, and Decidua;
RX MEDLINE=93126380; PubMed=7678348;
RA Paria B.C., Das S.K., Andrews G.K., Dey S.K.;
RT "Expression of the epidermal growth factor receptor gene is regulated
RT in mouse blastocysts during delayed implantation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RA Hibbs M.L.;
RT Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
RL -----
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=B6/C3; TISSUE=Liver;
RX MEDLINE=94170986; PubMed=8125255;
RA Luetke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,
RA Jenkins N.A., Lee D.C.;
RT "The mouse waved-2 phenotype results from a point mutation in the EGF
RT receptor tyrosine kinase.";
RL Genes Dev. 8:399-413(1994).
RN [5]
RP SEQUENCE OF 1-714 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91232866; PubMed=2030916;
RA Avivi A., Lax I., Ullrich A., Schlessinger J., Givol D., Morse B.;
RT "Comparison of EGF receptor sequences as a guide to study the ligand
RT binding site.";
RL Oncogene 6:673-676(1991).
RN [6]
RP SEQUENCE OF 969-1117 FROM N.A.
RC STRAIN=C3H;
RA Eisinger D.P., Serrero G.;
RT Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF, AND
CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- MISCELLANEOUS: BINDING OF EGF TO THE RECEPTOR LEADS TO
CC INTERNALIZATION OF THE EGF-RECEPTOR COMPLEX, INDUCTION OF THE
CC TYROSINE KINASE ACTIVITY, STIMULATION OF CELL DNA SYNTHESIS, AND
CC CELL PROLIFERATION.
CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X78987; CAA55587.1; -.
EMBL; U03425; AAA17899.1; -.
EMBL; X59698; CAA42219.1; -.
EMBL; L06864; AAA53029.1; -.
EMBL; Z12608; CAA78249.1; -.
HSSP; P11362; IFGI.
MGD; MGI:95294; Egfr.
InterPro: IPR000494; EGFR_L.
InterPro: IPR000719; Euk_pkinase.
InterPro: IPR002174; Furin-like.
InterPro: IPR001245; Tyr_kin.
Pfam; PF00757; Furin-like; 1.
Pfam; PF00069; pkinase; 1.
Pfam; PF01030; Recep_L_domain; 2.
SMART; SM00261; FU; 3.
SMART; SM00219; TyrKc; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
Transmembrane: Glycoprotein; Receptor; Signal; Transferase;
Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.
SIGNAL 1 24 POTENTIAL.
CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.
DOMAIN 25 647 EXTRACELLULAR (POTENTIAL).
TRANSMEM 648 670 POTENTIAL.
DOMAIN 671 1210 CYTOPLASMIC (POTENTIAL).
REPEAT 75 300 APPROXIMATE.
REPEAT 390 600 APPROXIMATE.
DOMAIN 1028 1071 SER-RICH.
DOMAIN 714 981 PROTEIN KINASE.
NP_BIND 720 728 ATP (BY SIMILARITY).
RL -----

RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,
RA O'Malley B.W.;
RT "Isolation of an evolutionarily conserved epidermal growth factor
RT receptor cDNA from human A431 carcinoma cells.";
RL Biochem. Biophys. Res. Commun. 124:125-132(1984).
[5]
RX SEQUENCE OF 1-29 FROM N.A.
RP MEDLINE=88217333; PubMed=3329716;
RA Haley J., Whittle N., Bennett P., Kinchington D., Ullrich A.,
RA Waterfield M.;
RT "The human EGF receptor gene: structure of the 110 kb locus and
RT identification of sequences regulating its transcription.";
RL Oncogene Res. 1:375-396(1987).
[6]
RX SEQUENCE OF 1-29 FROM N.A.
RP MEDLINE=91107677; PubMed=1988448;
RA Haley J.D., Waterfield M.D.;
RT "Contributory effects of de novo transcription and premature
RT transcript termination in the regulation of human epidermal growth
RT factor receptor proto-oncogene RNA synthesis.";
RL J. Biol. Chem. 266:1746-1753(1991).
[7]
RX SEQUENCE OF 1-29 FROM N.A.
RP MEDLINE=85270438; PubMed=2991899;
RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;
RT "Characterization and sequence of the promoter region of the human
RT epidermal growth factor receptor gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).
[8]
RX SEQUENCE OF 540.
RP Kohda D.;
RA Submitted (SEP-1997) to the SWISS-PROT data bank.
[9]
RX RECEPTOR ACTIVITY.
RP MEDLINE=84191554; PubMed=6325948;
RA Mroczkowski B., Mosig G., Cohen S.;
RT "ATP-stimulated interaction between epidermal growth factor receptor
RT and supercoiled DNA.";
RL Nature 309:270-273(1984).
[10]
RX PHOSPHORYLATION.
RP MEDLINE=89278137; PubMed=2543678;
RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,
RA Hawk R., Givol D., Ullrich A., Schlessinger J.;
RT "All autophosphorylation sites of epidermal growth factor (EGF)
RT receptor and HER2/neu are located in their carboxyl-terminal tails.
RT Identification of a novel site in EGF receptor.";
RL J. Biol. Chem. 264:10667-10671(1989).
[11]
RX REVIEW.
RP MEDLINE=87297456; PubMed=3039909;
RA Carpenter G.;
RT "Receptors for epidermal growth factor and other polypeptide
RT mitogens.";
RL Annu. Rev. Biochem. 56:881-914(1987).
CC -!- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
CC VACCINIA VIRUS GROWTH FACTOR.
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- MISCELLANEOUS: BINDING OF EGF TO THE RECEPTOR LEADS TO
CC INTERNALIZATION OF THE EGF-RECEPTOR COMPLEX, INDUCTION OF THE
CC TYROSINE KINASE ACTIVITY, STIMULATION OF CELL DNA SYNTHESIS, AND
CC CELL PROLIFERATION.
CC -!- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; X00588; CAA25240.1; -;
DR EMBL; X06370; CAA29668.1; -;
DR EMBL; X00663; CAA25282.1; -;
DR EMBL; M38425; AAA63171.1; -;
DR EMBL; M11234; AAA52370.1; -;
DR PIR; A00641; GQHUE;
DR PIR; A00642; GQHUE2;
DR PIR; A23062; A23062.
DR HSP; P11362; IFGI.
DR SWISS-2DPAGE; P00533; HUMAN.
DR MIM; 131550;
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR001245; Tyr_Kin.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Repeat; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 24
FT CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT DOMAIN 25 645 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 646 668 POTENTIAL.
FT DOMAIN 669 1210 CYTOPLASMIC (POTENTIAL).
FT REPEAT 75 300 APPROXIMATE.
FT REPEAT 390 600 APPROXIMATE.
FT DOMAIN 1025 1071 SER-RICH.
FT DOMAIN 712 979 PROTEIN KINASE.
FT NP_BIND 718 726 ATP (BY SIMILARITY).
FT BINDING 745 745 ATP (BY SIMILARITY).
FT ACT_SITE 837 837 BY SIMILARITY.
FT MOD_RES 678 678 PHOSPHORYLATION (BY PKC).
FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-).
FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-).
FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-).
FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 540 540 N -> K (IN REF. 1).
SQ SEQUENCE 1210 AA; 134277 MW; D8A2A50B4EFB6ED2 CRC64;

Query Match 20.8%; Score 62.5; DB 1; Length 1210;
Best Local Similarity 34.0%; Pred. No. 12;
Matches 17; Conservative 2; Mismatches 20; Indels 11; Gaps 3;
Oy 9 CSRGSSWADLKCMD-----CAS-----CRARPHSDFCLG-CAAAPAP 47
ID LMAI_HUMAN STANDARD; PRT; 3075 AA.
AC P25391;
DT 01-MAY-1992 (Rel. 22, Created)
Db 194 CPNGSCWAGEENCQKLTIIICAAQCSCRCRCGRKSPDCDCHNQCAAGCTG 243
RESULT 8
LMAI_HUMAN
ID LMAI_HUMAN STANDARD; PRT; 3075 AA.
AC P25391;
DT 01-MAY-1992 (Rel. 22, Created)

FT	DISULFID	744	757	BY SIMILARITY.
FT	DISULFID	760	769	BY SIMILARITY.
FT	DISULFID	772	788	BY SIMILARITY.
FT	DISULFID	791	806	BY SIMILARITY.
FT	DISULFID	793	816	BY SIMILARITY.
FT	DISULFID	819	828	BY SIMILARITY.
FT	DISULFID	831	846	BY SIMILARITY.
FT	DISULFID	849	863	BY SIMILARITY.
FT	DISULFID	851	870	BY SIMILARITY.
FT	DISULFID	873	882	BY SIMILARITY.
FT	DISULFID	885	899	BY SIMILARITY.
FT	DISULFID	902	914	BY SIMILARITY.
FT	DISULFID	904	921	BY SIMILARITY.
FT	DISULFID	923	932	BY SIMILARITY.
FT	DISULFID	935	948	BY SIMILARITY.
FT	DISULFID	951	963	BY SIMILARITY.
FT	DISULFID	953	969	BY SIMILARITY.
FT	DISULFID	971	980	BY SIMILARITY.
FT	DISULFID	983	995	BY SIMILARITY.
FT	DISULFID	998	1007	BY SIMILARITY.
FT	DISULFID	1000	1014	BY SIMILARITY.
FT	DISULFID	1016	1025	BY SIMILARITY.
FT	DISULFID	1028	1041	BY SIMILARITY.
FT	DISULFID	1044	1056	BY SIMILARITY.
FT	DISULFID	1046	1063	BY SIMILARITY.
FT	DISULFID	1065	1074	BY SIMILARITY.
FT	DISULFID	1077	1087	BY SIMILARITY.
FT	DISULFID	1403	1412	BY SIMILARITY.
FT	DISULFID	1405	1419	BY SIMILARITY.
FT	DISULFID	1422	1431	BY SIMILARITY.
FT	DISULFID	1434	1449	BY SIMILARITY.
FT	DISULFID	1452	1466	BY SIMILARITY.
FT	DISULFID	1454	1476	BY SIMILARITY.
FT	DISULFID	1479	1488	BY SIMILARITY.
FT	DISULFID	1491	1506	BY SIMILARITY.
FT	DISULFID	1509	1521	BY SIMILARITY.
FT	DISULFID	1511	1528	BY SIMILARITY.
FT	DISULFID	1530	1539	BY SIMILARITY.
FT	DISULFID	1542	1553	BY SIMILARITY.
FT	DISULFID	1556	1556	BY SIMILARITY.
FT	DISULFID	1560	1560	INTERCHAIN (PROBABLE).
FT	CARBOHYD	38	38	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	555	555	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	665	665	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	763	763	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	926	926	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	952	952	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1045	1045	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1407	1407	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1579	1579	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1596	1596	N-LINKED (GLCNAC. .) (POTENTIAL).
Query Match 20.8%; Score 62.5; DB 1; Length 3075;				
Best Local Similarity 23.5%; Pred. No. 26;				
Matches 16; Conservative 6; Mismatches 23; Indels 23; Gaps 2;				
Oy	4	PGTAPC-----SRGSSWSADLDCMDCASCARPHSDFC-----LGC 40		
Db	1395	PLVAPCPVCNNHSDTCDPNTGKLCMGCDNTAGDCHDVCISGYIKVGTGSASDCALCAC 1454		
Oy	41	AAAPPAPF 48		
Db	1455	PHSPPPASF 1462		
RESULT 9				
ID	FER_PEPAS	STANDARD;	PRT;	54 AA.
AC	P00193;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	FERREDOXIN.			

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X91245; CAA62627.1; -;
CC EMBL; AF090832; AAC61591.1; -;
CC EMBL; AE003672; AAF54063.1; -;
CC HSSP; Q05158; 1QJ1.
CC FlyBase; FBgn014863; Mlp84B.
CC InterPro; IPR001781; LIM.
CC Pfam; PF00412; LIM; 5.
CC ProDom; PD000094; LIM; 5.
CC SMART; SM00132; LIM; 5.
CC PROSITE; PS00478; LIM_DOMAIN_1; 5.
CC PROSITE; PS00023; LIM_DOMAIN_2; 5.
CC Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc; Myogenesis;
KW Developmental protein; Differentiation.
FT DOMAIN 12 63
FT DOMAIN 65 80
FT DOMAIN 66 71
FT DOMAIN 120 172
FT DOMAIN 175 180
FT DOMAIN 178 189
FT DOMAIN 222 274
FT DOMAIN 276 291
FT DOMAIN 325 377
FT DOMAIN 379 390
FT DOMAIN 421 473
FT DOMAIN 475 490
FT SEQUENCE 495 AA; 53525 MW; 2E559B9178E54C0E CRC64;

Query Match 20.4%; Score 61.5; DB 1; Length 495;
Best Local Similarity 31.1%; Pred. No. 7;
Matches 14; Conservative 4; Mismatches 20; Indels 7; Gaps 1;

QY 10 SRGSSWSADLRCMDCAS-----CRARPHSDFCLGCAAAPAP 47
DB 135 ARGSRWHKCFKCGTKCKGLDLSLCCAPDKNIYKGCYAKKFGP 179

RESULT 12

ID MT3_RAT STANDARD; PRT; 66 AA.
AC P37361;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALLOTHIONEIN-III (MT-III) (GROWTH INHIBITORY FACTOR) (GIF).
GN MT3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP MEDLINE=94018480; PubMed=8412560;
RA Kobayashi H., Uchida Y., Ihara Y., Nakajima K., Kohsaka S.,
RA Miyatake T., Tsuji S.;
RT "Molecular cloning of rat growth inhibitory factor cDNA and the
expression in the central nervous system."
RL Brain Res. Mol. Brain Res. 19:188-194(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Glial tumor;
RA Anourex M.C., Rethaus E., Wurch T., Colpaert F.C., Pauwels P.J.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Chapman G.A., Kille P.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: BINDS HEAVY METALS. CONTAINS ZINC AND COPPER ATOMS AND
CC ONLY A NEGLIGIBLE AMOUNT OF CADMIUM (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; S65838; AAB28366.1; -;
CC EMBL; X89603; CAA61762.1; -;
CC EMBL; Y08235; CAA69404.1; -;
CC HSSP; P18055; 1MRB.
CC InterPro; IPR003019; Metallthion.
CC InterPro; IPR000006; Metallthion_vert.
CC Pfam; PF00131; metalthio; 1.
CC PRINTS; PR00860; MTVERTEBRATE.
CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.

KW Metal-binding; Metal-thiolate cluster; Chelation; Zinc; Copper;
KW Acetylation.

FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT DOMAIN 1 30 BETA.
FT DOMAIN 31 66 ALPHA.
FT METAL 6 6 CLUSTER B (BY SIMILARITY).
FT METAL 8 8 CLUSTER B (BY SIMILARITY).
FT METAL 14 14 CLUSTER B (BY SIMILARITY).
FT METAL 16 16 CLUSTER B (BY SIMILARITY).
FT METAL 20 20 CLUSTER B (BY SIMILARITY).
FT METAL 22 22 CLUSTER B (BY SIMILARITY).
FT METAL 25 25 CLUSTER B (BY SIMILARITY).
FT METAL 27 27 CLUSTER B (BY SIMILARITY).
FT METAL 30 30 CLUSTER B (BY SIMILARITY).
FT METAL 34 34 CLUSTER A (BY SIMILARITY).
FT METAL 35 35 CLUSTER A (BY SIMILARITY).
FT METAL 37 37 CLUSTER A (BY SIMILARITY).
FT METAL 38 38 CLUSTER A (BY SIMILARITY).
FT METAL 42 42 CLUSTER A (BY SIMILARITY).
FT METAL 45 45 CLUSTER A (BY SIMILARITY).
FT METAL 49 49 CLUSTER A (BY SIMILARITY).
FT METAL 51 51 CLUSTER A (BY SIMILARITY).
FT METAL 62 62 CLUSTER A (BY SIMILARITY).
FT METAL 64 64 CLUSTER A (BY SIMILARITY).
FT METAL 65 65 CLUSTER A (BY SIMILARITY).
FT SEQUENCE 66 AA; 6809 MW; BE7538E8564EBF8 CRC64;

Query Match 20.3%; Score 61; DB 1; Length 66;
Best Local Similarity 35.0%; Pred. No. 1.4;
Matches 14; Conservative 2; Mismatches 22; Indels 2; Gaps 2;

QY 4 PCTAPCSRGSSWSADLRCMDCASCARPHSDFCLGCAAA 43
DB 3 PETCPCTGGCTCS-DKC-KCKGCKTCKKSCSCCPA 40

RESULT 13

ID MT3_MOUSE STANDARD; PRT; 68 AA.
AC P28184;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN-III (MT-III) (GROWTH INHIBITORY FACTOR) (GIF).
GN MT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92335292; PubMed=1631128;
 RA Palmer R.D., Findley S.D., Whitmore T.E., Durnam D.M.;
 RT "MT-III, a brain-specific member of the metallothionein gene family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6333-6337(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94314230; PubMed=8039715;
 RA Naruse S., Igarashi S., Furuya T., Kobayashi H., Miyatake T.,
 RA Tsuji S.;
 RT "Structures of the human and mouse growth inhibitory factor-encoding
 genes.";
 RL Gene 144:283-287(1994).
 CC -!- FUNCTION: INHIBITS SURVIVAL AND NEURITE FORMATION OF CORTICAL
 CC NEURONS IN VITRO.
 CC -!- FUNCTION: BINDS HEAVY METALS. CONTAINS ZINC AND COPPER ATOMS AND
 CC ONLY A NEGLIGIBLE AMOUNT OF CADMIUM (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M93310; AAA39529.1; -;
 DR EMBL; S72046; AAB31397.1; -;
 DR PIR; A46034; A46034.
 DR HSSP; P18055; 1MRB.
 DR MGP; MGI:97173; Mt3.
 DR InterPro; IPR003019; Metallothion.
 DR InterPro; IPR000006; Metallothion_vert.
 DR Pfam; PF00131; metalthio; 1.
 DR PRINTS; PR00860; MTVERTEBRATE.
 DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
 KW Metal-binding; Metal-thiolate cluster; Chelation; Zinc; Copper;
 KW Acetylation.
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 FT DOMAIN 1 30 BETA.
 FT DOMAIN 31 60 ALPHA.
 FT METAL 6 6 CLUSTER B (BY SIMILARITY).
 FT METAL 8 8 CLUSTER B (BY SIMILARITY).
 FT METAL 14 14 CLUSTER B (BY SIMILARITY).
 FT METAL 16 16 CLUSTER B (BY SIMILARITY).
 FT METAL 20 20 CLUSTER B (BY SIMILARITY).
 FT METAL 22 22 CLUSTER B (BY SIMILARITY).
 FT METAL 25 25 CLUSTER B (BY SIMILARITY).
 FT METAL 27 27 CLUSTER B (BY SIMILARITY).
 FT METAL 30 30 CLUSTER B (BY SIMILARITY).
 FT METAL 34 34 CLUSTER A (BY SIMILARITY).
 FT METAL 35 35 CLUSTER A (BY SIMILARITY).
 FT METAL 37 37 CLUSTER A (BY SIMILARITY).
 FT METAL 38 38 CLUSTER A (BY SIMILARITY).
 FT METAL 42 42 CLUSTER A (BY SIMILARITY).
 FT METAL 45 45 CLUSTER A (BY SIMILARITY).
 FT METAL 49 49 CLUSTER A (BY SIMILARITY).
 FT METAL 51 51 CLUSTER A (BY SIMILARITY).
 FT METAL 64 64 CLUSTER A (BY SIMILARITY).
 FT METAL 66 66 CLUSTER A (BY SIMILARITY).
 FT METAL 67 67 CLUSTER A (BY SIMILARITY).
 FT SEQUENCE 68 AA; 7009 MW; 791AF60E38FED3CA CRC64;
 Query Match 20.3%; Score 61; DB 1; Length 68;
 Best Local Similarity 35.0%; Pred. No. 1.5;
 Matches 14; Conservative 2; Mismatches 22; Indels 2; Gaps 2;
 Oy 4 PGTAPCSRGWSADLDKCMDCASCARPHSDFCIGCAA 43
 | | | | | | | | | | | | | | | | | | | | | |

Db 3 PETCPPTGGSCS-DKC-KCKGCKTCKKSCSCCPA 40
 RESULT 14
 ID LDLR_RABIT STANDARD; PRT; 837 AA.
 AC P20063;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE LOW-DENSITY LIPOPROTEIN RECEPTOR (LDL RECEPTOR) (FRAGMENT).
 DL LDLR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86208167; PubMed=3010466;
 RA Yamamoto T., Bishop R.W., Brown M.S., Goldstein J.L., Russell D.W.;
 RT "Deletion in cysteine-rich region of LDL receptor impedes transport
 RT to cell surface in WHHL rabbit.";
 RL Science 232:1230-1237(1986).
 CC -!- FUNCTION: BINDS LDL, THE MAJOR CHOLESTEROL-CARRYING LIPOPROTEIN
 CC OF PLASMA, AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS. IN ORDER
 CC TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST FIRST
 CC CLUSTER INTO CLATHRIN-COATED PITS.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: CONTAINS 7 LDL-RECEPTOR CLASS A DOMAINS.
 CC -!- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M11501; AAA31383.1; -;
 DR PIR; A29512; A29512.
 DR HSSP; P01130; 1LDL.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR002172; LDL_recept_A.
 DR InterPro; IPR000033; Ldl_rcptor_rep.
 DR Pfam; PF00057; ldl_recept_a; 7.
 DR Pfam; PF00058; ldl_recept_b; 5.
 DR SMART; SM00179; EGF_CA; 1.
 DR SMART; SM00001; EGF_like; 2.
 DR SMART; SM00192; LDLA; 7.
 DR SMART; SM00135; LY; 5.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS01209; LDLA_1; 7.
 DR PROSITE; PS00068; LDLA_2; 7.
 KW Glycoprotein; LDL; Cholesterol metabolism; Lipid transport;
 KW Endocytosis; Coated pits; Transmembrane; Receptor; EGF-like domain;
 KW Repeat.
 FT NON_TER 1 1
 FT DOMAIN <1 765
 FT TRANSMEM 766 787
 FT DOMAIN 788 837
 FT DOMAIN 12 52
 FT DOMAIN 53 93
 FT DOMAIN 94 132
 FT DOMAIN 133 173
 FT DOMAIN 182 220
 FT DOMAIN 221 259
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 LDL-RECEPTOR CLASS A 1.
 LDL-RECEPTOR CLASS A 2.
 LDL-RECEPTOR CLASS A 3.
 LDL-RECEPTOR CLASS A 4.
 LDL-RECEPTOR CLASS A 5.
 LDL-RECEPTOR CLASS A 6.

FT DOMAIN 261 300 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 301 340 EGF-LIKE 1.
FT REPEAT 341 380 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 384 425 LDL-RECEPTOR CLASS B 1.
FT REPEAT 426 471 LDL-RECEPTOR CLASS B 2.
FT REPEAT 473 515 LDL-RECEPTOR CLASS B 3.
FT REPEAT 516 559 LDL-RECEPTOR CLASS B 4.
FT REPEAT 560 602 LDL-RECEPTOR CLASS B 5.
FT REPEAT 603 644 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 650 699 EGF-LIKE 3.
FT DOMAIN 700 749 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 14 26 BY SIMILARITY.
FT DISULFID 21 39 BY SIMILARITY.
FT DISULFID 33 50 BY SIMILARITY.
FT DISULFID 55 69 BY SIMILARITY.
FT DISULFID 62 82 BY SIMILARITY.
FT DISULFID 76 91 BY SIMILARITY.
FT DISULFID 96 108 BY SIMILARITY.
FT DISULFID 103 121 BY SIMILARITY.
FT DISULFID 115 130 BY SIMILARITY.
FT DISULFID 135 147 BY SIMILARITY.
FT DISULFID 142 160 BY SIMILARITY.
FT DISULFID 154 171 BY SIMILARITY.
FT DISULFID 184 196 BY SIMILARITY.
FT DISULFID 191 209 BY SIMILARITY.
FT DISULFID 203 218 BY SIMILARITY.
FT DISULFID 223 235 BY SIMILARITY.
FT DISULFID 230 248 BY SIMILARITY.
FT DISULFID 242 257 BY SIMILARITY.
FT DISULFID 262 276 BY SIMILARITY.
FT DISULFID 271 289 BY SIMILARITY.
FT DISULFID 283 300 BY SIMILARITY.
FT DISULFID 305 316 BY SIMILARITY.
FT DISULFID 312 325 BY SIMILARITY.
FT DISULFID 327 339 BY SIMILARITY.
FT DISULFID 345 355 BY SIMILARITY.
FT DISULFID 351 364 BY SIMILARITY.
FT DISULFID 366 379 BY SIMILARITY.
FT DISULFID 654 668 BY SIMILARITY.
FT DISULFID 664 683 BY SIMILARITY.
FT DISULFID 685 698 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 644 644 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 837 AA; 91406 MW; ED8D231E234400A9 CRC64;

Query Match 19.9%; Score 60; DB 1; Length 837;
Best Local Similarity 36.5%; Pred. No. 16;
Matches 19; Conservative 1; Mismatches 14; Indels 18; Gaps 4;
OY 4 PGTAPCR-----GS-----SWSADLD-KC-----MDCASCARPHSDFC 37
DB 179 PGRGPCSRHFHCGEGCVHSMRCDGDADCRDGSDDRDCAATCRPDEFQC 230
ID PKCS_MOUSE STANDARD; PRT; 1877 AA.
RESULT 15
PKCS_MOUSE
AC Q04592; Q62040;
DT 01-FEB-1995 (Rel. 31, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROTEIN CONVERTASE SUBTILISIN/KEXIN TYPE 5 PRECURSOR (EC 3.4.21.-)
DE (PROTEIN CONVERTASE PC5) (SUBTILISIN/KEXIN-LIKE PROTEASE PC5)
DE (CONVERTASE PC5) (PC6) (SUBTILISIN-LIKE PROTEIN CONVERTASE 6)
GN (SPC6).
DE PCSK5.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PC5B).
RC STRAIN=ICR; TISSUE=Intestine;
RX MEDLINE=93327934; PubMed=8335106;
RA Nakagawa T., Murakami K., Nakayama K.;
RT "Identification of an isoform with an extremely large Cys-rich region
of PC6, a Kex2-like processing endoprotease.";
RN FEBS Lett. 327:165-171(1993).
RP SEQUENCE FROM N.A. (ISOFORM PC5A).
RC TISSUE=Brain, and Intestine;
RX MEDLINE=93224489; PubMed=8468318;
RA Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
RA Nakayama K.;
RT "Identification and functional expression of a new member of the
mammalian Kex2-like processing endoprotease family: its striking
structural similarity to PACE4.";
RL J. Biochem. 113:132-135(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM PC5A).
RC TISSUE=Adrenal cortex;
RX MEDLINE=93342056; PubMed=8341687;
RA Lussan J., Vieau D., Hamelin J., Day R., Chretien M., Seidah N.G.;
RT "cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a
candidate proprotein convertase expressed in endocrine and
nonendocrine cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
RN [4]
RP PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.
RX MEDLINE=97103178; PubMed=8947550;
RA De Bie I., Marcinkiewicz M., Malide D., Lazure C., Nakayama K.,
RA Bendayan M., Seidah N.G.;
RT "The isoforms of proprotein convertase PC5 are sorted to different
subcellular compartments.";
RL J. Cell Biol. 135:1261-1275(1996).
RN [5]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=96293359; PubMed=8698813;
RA Constam D.B., Calton M., Robertson E.J.;
RT "SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone
morphogenetic proteins at distinct sites during embryogenesis.";
RL J. Cell Biol. 134:181-191(1996).
RN [6]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=97436919; PubMed=9291583;
RA Rancourt S.L., Rancourt D.E.;
RT "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
implantation, somitogenesis, and skeletal formation.";
RL Dev. Genet. 21:75-81(1997).
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE
FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
OF GROWTH FACTORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
PROTEIN BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: PC5A IS SECRETED THROUGH THE REGULATED
SECRETORY PATHWAY. PC5B IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
EARLY ENDOSOMES.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; PC5B/LONG (SHOWN HERE)
AND PC5A/SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PC5A IS EXPRESSED IN MOST TISSUES BUT IS MOST
ABUNDANT IN THE INTESTINE AND ADRENALS. PC5B IS EXPRESSED IN THE
INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,
BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT
E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA. AT
E7.5, INTENSE EXPRESSION IN EXTRAEMBRYONIC ENDODERM, AMNION AND
NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK
SAC FOLLOWED BY A CONFINATION TO DERMAMYOTOME COMPARTMENT. BETWEEN

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2002, 15:23:24 ; Search time 54.85 Seconds
(without alignments)
138.672 Million cell updates/sec

Title: US-09-742-454a-7_COPY_28_79

Perfect score: 301

Sequence: 1 EQAGPAPCRSGSSWADLD.....HSDFCLGCAAPAPFRLW 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	301	100.0	129	4 Q9NP84	Q9np84 homo sapien
2	276	91.7	129	11 Q9QW3	Q9qzw3 mus musculu
3	276	91.7	129	11 Q9CR75	Q9cr75 mus musculu
4	67.5	22.4	160	2 Q86474	Q86474 rhodobacter
5	67	22.3	94	4 Q9HCS0	Q9hcs0 homo sapien
6	66.5	22.1	420	5 P91776	P91776 pacifastacu
7	66	21.9	549	2 Q9HVT3	Q9hvt3 pseudomonas
8	64.5	21.4	112	11 Q9J148	Q9j148 mus musculu
9	64.5	21.4	768	13 Q9ZT48	Q9zt48 cyprinus ca
10	64	21.3	115	4 Q9NZF1	Q9nzf1 homo sapien
11	64	21.3	473	11 Q9ESR0	Q9ese0 rattus norv
12	64	21.3	643	11 Q9ERV6	Q9erv6 mus musculu
13	64	21.3	655	11 Q9WVF5	Q9wvf5 mus musculu
14	64	21.3	669	4 Q75441	Q75441 homo sapien
15	64	21.3	1210	11 Q9EP98	Q9ep98 mus musculu
16	64	21.3	1511	4 Q75412	Q75412 homo sapien
17	64	21.3	1587	4 Q00508	Q00508 homo sapien
18	63.5	21.1	85	1 Q9P9E6	Q9p9e6 methanococc
19	63.5	21.1	708	13 P87363	P87363 gallus gall

ALIGNMENTS

RESULT 1

Q9NP84	Q9NP84	PRELIMINARY;	PRT;	129 AA.
AC	Q9NP84;			
DT	01-OCT-2000 (Tremblrel. 15, Created)			
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)			
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)			
DE	TYPE I TRANSMEMBRANE PROTEIN PRECURSOR (TYPE I TRANSMEMBRANE PROTEIN			
DE	FN14).			
GN	FN14.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Tanaka S., Sugimachi K.;			
RT	"Human homologue of Fn14.";			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PLACENTA;			
RX	MEDLINE=20216634; PubMed=10751351;			
RA	Feng S.-L.Y., Guo Y., Factor V.M., Thorgeirsson S.S., Bell D.W.,			
RT	Testa J.R., Peiffley K.A., Winkles J.A.;			
RT	"The Fn14 immediate-early response gene is induced during liver			
RT	regeneration and highly expressed in both human and murine			
RT	hepatocellular carcinoma.";			
RL	Am. J. Pathol. 156:1253-1261(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=ENDOMETRIAL ADENOCARCINOMA;			
RA	Strausberg R.;			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB035480; BAA94792.1;			
DR	EMBL; AF191148; AAF69108.1;			
DR	EMBL; BC002718; AAH02718.1;			
DR	Signal; Transmembrane.			
KW	POTENTIAL.			
FT	SIGNAL 9 31			
SQ	SEQUENCE 129 AA; 13911 MW; BF3DFB9C1E1C448 CRC64;			

Q92795 homo sapien
Q9fca1 streptomyce
Q9h2c9 homo sapien
Q14226 homo sapien
Q9gzx1 homo sapien
P91404 caenorhabdi
O27680 methanobact
O24155 nicotiana t
Q91jg5 mus musculu
Q19088 caenorhabdi
P92163 strongyloce
O75136 homo sapien
Q9qx70 rattus norv
O75197 homo sapien
Q9up66 homo sapien
Q9ues7 homo sapien
O75851 homo sapien
Q9daw5 mus musculu
Q9qzf2 mus musculu
Q913j0 pseudomonas
Q83022 rhodobacter
Q9rad6 rhodobacter
Q95054 homo sapien
Q9qee6 indian citr
Q95407 homo sapien
Q9uja2 homo sapien

```
Query Match      100.0%; Score 301; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.2e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EQAGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLLW 52
Db 28 EQAGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLLW 79

RESULT 2
Q90ZW3 ID Q90ZW3 PRELIMINARY; PRT; 129 AA.
AC Q90ZW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TYPE I TRANSMEMBRANE PROTEIN FN14.
GN FGFRP2 OR FN14.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C;
RX MEDLINE=20020297; PubMed=10551889;
RA Meighan-Mantha R.L., Hsu D.K.W., Guo Y., Brown S.A.N., Feng S.-L.Y.,
RA Peifley K.A., Alberts G.F., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Richards C.M., Winkles J.A.;
RT "The mitogen-inducible Fn14 gene encodes a type I transmembrane
RL protein that modulates fibroblast adhesion and migration.";
RL J. Biol. Chem. 274:33166-33176(1999).
DR EMBL: AF156164; AAF07882.1; -
DR MGD: MGI:1351484; Fgfrp2.
KW Transmembrane.
SQ SEQUENCE 129 AA; 13637 MW; 14B5C68EEF493385 CRC64;

Query Match      91.7%; Score 276; DB 11; Length 129;
Best Local Similarity 92.3%; Pred. No. 2.8e-25;
Matches 48; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EQAGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLLW 52
Db 28 EQAGTSPCSGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLLW 79

RESULT 3
Q9CR75 ID Q9CR75 PRELIMINARY; PRT; 129 AA.
AC Q9CR75;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FIBROBLAST GROWTH FACTOR REGULATED PROTEIN 2.
GN FGFRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA, AND CEREBELLUM;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo I., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
```

```
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK005530; BAB24101.1; -
DR EMBL: AK005382; BAB23989.1; -
DR MGD: MGI:1351484; Fgfrp2.
SQ SEQUENCE 129 AA; 13641 MW; 1665C68B4D9A9253 CRC64;

Query Match      91.7%; Score 276; DB 11; Length 129;
Best Local Similarity 92.3%; Pred. No. 2.8e-25;
Matches 48; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EQAGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLLW 52
Db 28 EQAGTSPCSGSSWSADLDKCMDCASCARPHSDFCGLGCAAAAPPAPFRLLW 79

RESULT 4
O86474 ID O86474 PRELIMINARY; PRT; 160 AA.
AC O86474;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE NAF.
GN NAF.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 158;
RX MEDLINE=96332666; PubMed=8730872;
RA Reyes F., Roldan D., Klipp W., Castillo F., Moreno-Vivian C.;
RT "Isolation of periplasmic nitrate reductase genes from Rhodobacter
RT sphaeroides DSM 158: structural and functional differences among
RT prokaryotic nitrate reductases.";
RL Mol. Microbiol. 19:1307-1318(1996).
DR EMBL: Z46806; CAA86832.1; -
DR HSSP: P00195; 1CLF.
DR InterPro: IPR001450; 4FE4S_ferredoxin.
DR Pfam: PF00037; fer4; 3.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 3.
KW Iron-sulfur.
SQ SEQUENCE 160 AA; 16392 MW; 01FCADE2EACB9EC6 CRC64;

Query Match      22.4%; Score 67.5; DB 2; Length 160;
Best Local Similarity 29.7%; Pred. No. 1;
Matches 19; Conservative 4; Mismatches 18; Indels 23; Gaps 2;

QY 3 APGTAPCSRGSSWSADLDKCMDCASCARPHSDFCGLGCAAA 43
Db 18 APTRPWTTR----EADNARTGCAACACACGACGIVRMEAGLPQIAFAGTCSCFGCAAEA 73

QY 44 PPAP 47
Db 74 CPAP 77

RESULT 5
Q9HCS0 ID Q9HCS0 PRELIMINARY; PRT; 94 AA.
```

AC Q9HGS0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TYPE I TRANSMEMBRANE PROTEIN PRECURSOR.
GN FN14 VARIANT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanaka S., Sugimachi K.;
RT "Human homologue of Fn14.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB035481; BAB17850.1; -
KW Signal; Transmembrane.
FT SIGNAL 9 31
FT SIGNAL 9 31
SQ SEQUENCE 94 AA; 10191 MW; E3C52B1308DC768F CRC64;

Query Match 22.3%; Score 67; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 AAAPPAPFRLW 52
DB 33 AAAPPAPFRLW 44

P91776
ID P91776 PRELIMINARY; PRT; 420 AA.
AC P91776;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PACIFASTIN LIGHT CHAIN PRECURSOR.
OS Pacifastacus leniusculus (Signal crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Astacidea; Astacidae; Pacifastacus.
OX NCBI_TaxID=6720;
RN [1]
RP SEQUENCE FROM N.A.
RA Liang Z., Sottrup-Jensen L., Aspan A., Hall M., Soderhall K.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81825; AAC64661.1; -
DR HSP; P80060; 1PMC.
DR InterPro: IPR001007; VWF.
DR SMART; SM00011; VWC_def; 2.
DR SMART; SM00215; VWC_out; 1.
KW Signal.
FT SIGNAL 1 15
FT CHAIN 16 420
FT CHAIN 16 420
SQ SEQUENCE 420 AA; 45418 MW; 037250AFA5C6FBC0 CRC64;

Query Match 22.1%; Score 66.5; DB 5; Length 420;
Best Local Similarity 31.4%; Pred. No. 2.8;
Matches 16; Conservative 5; Mismatches 17; Indels 13; Gaps 2;

QY 1 EQAPGTAPCSRGSSWADLDKCM---DCASCRAR-----PHSDFCL 38
DB 264 EPRGGEAVCTDGSKWKDDCNWCTCNNGSASCTEKLCOYKPDGSLPNDMVCV 314

P91773
ID P91773 PRELIMINARY; PRT; 549 AA.
AC Q9HVT3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN PA4488.
GN PA4488.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01.
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Miziouchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004862; AAG07876.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 549 AA; 61710 MW; DAB72175DB036B7C CRC64;

Query Match 21.9%; Score 66; DB 2; Length 549;
Best Local Similarity 36.7%; Pred. No. 3.9;
Matches 18; Conservative 3; Mismatches 16; Indels 12; Gaps 2;

QY 4 PGTAPCSRGSSWADLDKCMDCASCARPHSDFCGLCAAPAPFRLW 52
DB 380 PASAPTRRTAAWSADL---VLGATVNVHSD-----QPGPARLSW 416

P9148
ID P9148 PRELIMINARY; PRT; 112 AA.
AC Q9J148;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ONZIN (10 DAY OLD MALE PANCREAS CDNA, RIKEN FULL-LENGTH ENRICHED
DE LIBRARY, CLONE:181008C09, FULL INSERT SEQUENCE).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherwin J.R.A., Sharkey A.M., Smith S.K.;
RT "Identification of LIF regulated genes in the mouse uterus.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

```
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AF263458; AAF76887.1; -.
DR EMBL; AK007369; BAB24991.1; -.
SQ SEQUENCE 112 AA; 12353 MW; 4FFBD5E8D91F0DA8 CRC64;

Query Match 21.4%; Score 64.5; DB 11; Length 112;
Best Local Similarity 34.18; Pred. No. 1.7;
Matches 14; Conservative 2; Mismatches 24; Indels 1; Gaps 1;

Qy 4 PGTAPCSGSSWSADLCKM-DCASCRARPHSDFCGCAA 43
. || : || | | | | | | | | | | | | | | | |
Db 12 PGFVRAPQNSNWTSLCDGVCGLGTCFTCLGCOVA 52

RESULT 9
Q98TH8 PRELIMINARY; PRT; 768 AA.
AC Q98TH8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE INTEGRIN BETA2-CHAIN.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura M., Fujiki K., Nakao M., Shin D., Yano T.;
RT "cDNA cloning of a carp homologue of mammalian integrin beta2-chain.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB013074; BAB39130.1; -.
SQ SEQUENCE 768 AA; 84402 MW; 82564A9AA3ABDD65 CRC64;

Query Match 21.4%; Score 64.5; DB 13; Length 768;
Best Local Similarity 40.58; Pred. No. 7.5;
Matches 17; Conservative 4; Mismatches 14; Indels 7; Gaps 3;

Qy 9 CSRGSSWSADLCKMDCASCRARP--HSDFCGCAAAPAPF 48
. || | : || | | | | | | | | | | | | | | |
Db 600 CERGYSGA----KCKTSSCQ-RPKESGSCVEIAFGTGP 636

RESULT 10
Q9NZF1 PRELIMINARY; PRT; 115 AA.
AC Q9NZF1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE BM-004.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhao M., Song H., Li N., Peng Y., Han Z., Chen Z.;
RT "A novel gene expressed in human bone marrow.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208846; AAF64260.1; -.
SQ SEQUENCE 115 AA; 12507 MW; 24B6D7940390C6BF CRC64;

Query Match 21.3%; Score 64; DB 4; Length 115;
Best Local Similarity 30.88; Pred. No. 2;
Matches 16; Conservative 3; Mismatches 23; Indels 10; Gaps 2;
```

```
Qy 2 QAP-----GTAPCSGSSWSADLCKM-DCASCRARPHSDFCGCAA 43
. || | : || | | | | | | | | | | | | | | |
Db 4 QAPVVVVTPQGVGPAPQNSNWTGMDCFSDGVCGLGTCFTCLGCOVA 55

RESULT 11
Q9ESE0 PRELIMINARY; PRT; 473 AA.
AC Q9ESE0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR RELATED PROTEIN.
GN ERFP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-GASTRO-DUODENAL MUCOUS;
RA Yu Y., Moshier J.A., Majumdar A.P.N.;
RT "Cloning of a novel EGFR-related peptide: A putative negative
RT regulator of EGFR.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187818; AAG17037.1; -.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR002174; Furin-like.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR SMART; SM00261; FU; 1.
KW Receptor.
SQ SEQUENCE 473 AA; 52903 MW; 4E0D7C3074F28973 CRC64;

Query Match 21.3%; Score 64; DB 11; Length 473;
Best Local Similarity 31.7%; Pred. No. 5.9;
Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

Qy 4 PGTAP-----CSRGSSWSADLCKMD-----CAS-----CRARPHSDFCG-CAAPAP 47
. || | : || | | | | | | | | | | | | | | |
Db 184. PSSCPKCDPSPGNGSCGGGEEGCKLTIIICAQQCHRCRGRSPSDCHNQCAAGCTGP 243

RESULT 12
Q9ERV6 PRELIMINARY; PRT; 643 AA.
AC Q9ERV6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR ISOFORM 2.
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C3H/101, 129/SVJ, 129/SVEVTA;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Mahle N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275366; AAG28046.1; -.
DR EMBL; AF275364; AAG28046.1; JOINED.
DR EMBL; AF275365; AAG28046.1; JOINED.
DR MGD; MGI:95294; Egfr.
```

```

DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR000494; EGFR_L.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recept_L_domain; 2.
DR SMART: SM00261; Recept_L.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 643 AA; 71477 MW; DEF22002C84911B1 CRC64;

Query Match 21.3%; Score 64; DB 11; Length 643;
Best Local Similarity 31.7%; Pred. No. 7.5;
Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

QY 4 PGTPAP-----CSRGSWSADLDKCMD-----CAS-----CRARPHSDFCLG-CAAAPAP 47
DB 184 PSSCPKDPSPNGSCWGGGEENCQKLTKIIAQQSHRCGRSPSDCHNQCAAGCTGP 243

RESULT 13
Q9WVF5 PRELIMINARY; PRT; 655 AA.
AC Q9WVF5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE EPIDERMAL GROWTH FACTOR RECEPTOR (EPIDERMAL GROWTH FACTOR RECEPTOR
DE ISOFORM 3).
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RA Reiter J.L., Threadgill D.W., Danielsen A.J., Schell C.,
RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
RA Maible N.J.;
RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode
RT Carboxy-terminal Truncated Receptors.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV/TAC;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maible N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=5781/6J; TISSUE=LIVER;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

```

```

RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AF124513; AAD44149.1; -.
DR EMBL: AF275366; AAG28047.1; -.
DR EMBL: AF275364; AAG28047.1; JOINED.
DR EMBL: AF275365; AAG28047.1; JOINED.
DR EMBL: AK004944; BAB23688.1; -.
DR EMBL: AK004883; BAB23641.1; -.
DR EMBL: AK004911; BAB23662.1; -.
DR MGD: MGI:95294; Egrf.
DR InterPro: IPR000494; EGFR_L.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recept_L_domain; 2.
DR SMART: SM00261; Recept_L.
KW Receptor.
SQ SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;

Query Match 21.3%; Score 64; DB 11; Length 655;
Best Local Similarity 31.7%; Pred. No. 7.6;
Matches 19; Conservative 2; Mismatches 23; Indels 16; Gaps 4;

QY 4 PGTPAP-----CSRGSWSADLDKCMD-----CAS-----CRARPHSDFCLG-CAAAPAP 47
DB 184 PSSCPKDPSPNGSCWGGGEENCQKLTKIIAQQSHRCGRSPSDCHNQCAAGCTGP 243

RESULT 14
O75441 PRELIMINARY; PRT; 669 AA.
AC O75441;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LATENT TRANSFORMING GROWTH FACTOR-BETA BINDING PROTEIN 4 (FRAGMENT).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98325059; PubMed=9660815;
RA Saharinen J., Taipale J., Monni O., Keski-Oja J.;
RT "Identification and characterization of a new latent transforming
RT growth factor-beta-binding protein, LTBP-4.";
RL J. Biol. Chem. 273:18459-18469(1998).
DR EMBL: AF054502; AAC39882.1; -.
DR HSP: P35555; 1EMN.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002212; TB.
DR Pfam: PF00008; EGF; 6.
DR Pfam: PF00683; TB; 2.
DR SMART: SM00179; EGF_CA; 6.
DR SMART: SM00001; EGF_Like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 6.
DR PROSITE: PS01186; EGF_2; UNKNOWN_2.
DR PROSITE: PS01187; EGF_CA; 8.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1
FT NON_TER 669
SQ SEQUENCE 669 AA; 71701 MW; 2C76C44AB11428A4 CRC64;

Query Match 21.3%; Score 64; DB 4; Length 669;
Best Local Similarity 35.6%; Pred. No. 7.7;
Matches 16; Conservative 3; Mismatches 14; Indels 12; Gaps 2;

```

